Submitted to the Annals of Applied Probability
arXiv: math.PR/0909.4790

ERROR ANALYSIS OF TAU-LEAP SIMULATION METHODS

BY DAVID F. ANDERSON * ARNAB GANGULY * AND THOMAS G. KURTZ *

*University of Wisconsin - Madison**

We perform an error analysis for numerical approximation methods of continuous time Markov chain models commonly found in the chemistry and biochemistry literature. The motivation for the analysis is to be able to compare the accuracy of different approximation methods and, specifically, Euler tau-leaping and midpoint tau-leaping. We perform our analysis under a scaling in which the size of the time discretization is inversely proportional to some (bounded) power of the norm of the state of the system. We argue that this is a more appropriate scaling than that found in previous error analyses in which the size of the time discretization goes to zero independent of the rest of the model. Under the present scaling we show that midpoint tau-leaping achieves a higher order of accuracy, in both a weak and a strong sense, than Euler tau-leaping; a result that is in contrast to previous analyses. We present examples that demonstrate our findings.

1. Introduction. This paper provides an error analysis for numerical approximation methods for continuous time Markov chain models that are becoming increasingly common in the chemistry and biochemistry literature. Our goals of the paper are two-fold. First, we want to demonstrate the importance of considering appropriate scalings in which to carry out error analyses for the methods of interest. Second, we wish to provide such an error analysis in order to compare the accuracy of two different approximation methods. We perform our analysis on the Euler tau-leaping method first presented in [11] and a midpoint tau-leaping method developed below, which is only a slight variant of one presented in [11]. The midpoint tau-leaping method will be demonstrated to be more accurate than Euler tau-leaping in both a strong and a weak sense, a result that is in contrast to previous error analyses. We will discuss why previous error analyses made differing predictions than does ours and argue that the scaling provided here, or variants thereof, is a more natural and appropriate choice for error analyses of such methods. We also provide examples that demonstrate our findings.

1.1. *The basic model.* The motivation for the class of mathematical models under consideration comes from chemistry and biochemistry, and more generally from population processes (though we choose the language of chemistry throughout the paper). We assume the existence of a chemical reaction system consisting of (i) d chemical species $\{S_1, S_2, \ldots, S_d\}$ and (ii) a finite set of possible reactions, which we index by k. Each reaction requires some number of the species as inputs and provides some number of the species as outputs. For example, the reaction $S_1 \to 2S_2$ would require one molecule of S_1 for the input and provide two molecules of S_2 for the output. If reaction k occurs at time t, then the state of the system $X(t) \in \mathbb{Z}^d_{\geq 0}$ is updated via addition of the *reaction vector* $\nu_k \in \mathbb{Z}^d$, which represents the net change in the abundances of the underlying species:

$$X(t) = X(t-) + \nu_k.$$

Returning briefly to the example $S_1 \to 2S_2$, the associated reaction vector for this reaction would be $[-1, 2, 0, \ldots, 0]^T$. Finally, we denote by ν_k^s the vector in $\mathbb{Z}^d_{\geq 0}$ representing the source of the kth reaction. Returning again to the example $S_1 \to 2S_2$, the source vector for this reaction is $\nu_k^s = [1, 0, \ldots, 0]^T$.

*Grant support from NSF grant DMS-0553687.
AMS 2000 subject classifications: Primary 60H35, 65C99; Secondary 92C40
Keywords and phrases: tau-leaping, simulation, error analysis, reaction networks, Markov chain, chemical master equation

We assume that the waiting times for the k reactions are exponentially distributed with intensity functions $\lambda_k : \mathbb{R}^d_{\geq 0} \to \mathbb{R}_{\geq 0}$. We extend each λ_k to all of $\mathbb{R}^d$ by setting it to zero outside $\mathbb{R}^d_{\geq 0}$. This model is a continuous time Markov chain in $\mathbb{Z}^d_{\geq 0}$ with generator

$$(\mathcal{A}f)(x) = \sum_k \lambda_k(x)(f(x+\nu_k) - f(x)), \tag{1.1}$$

where $f : \mathbb{Z}^d \to \mathbb{R}$ is arbitrary. Kolmogorov's forward equation for this model, termed the "chemical master equation" in the chemistry and biology literature, is

$$\frac{d}{dt}P(x,t|\pi) = \sum_k P(x-\nu_k,t|\pi)\lambda_k(x-\nu_k) - \sum_k P(x,t|\pi)\lambda_k(x),$$

where for $x \in \mathbb{Z}^d_{\geq 0}$ $P(x,t|\pi)$ represents the probability that $X(t) = x$, conditioned upon the initial distribution π. One representation for path-wise solutions to this model uses a random time change of Poisson processes:

$$X(t) = X(0) + \sum_k Y_k\left(\int_0^t \lambda_k(X(s))ds\right)\nu_k, \tag{1.2}$$

where the Y_k are independent, unit-rate Poisson processes (see, for example, [16]). Note that $\tilde{X}(t) \stackrel{\text{def}}{=} X(t) - \sum_k \int_0^t \lambda_k(X(s))ds\,\nu_k$ is a martingale with quadratic covariation matrix $[X]_t = \sum_k Y_k\left(\int_0^t \lambda_k(X(s))ds\right)\nu_k\nu_k^T$.

A common choice of intensity function for chemical reaction systems, and the one we adopt throughout, is mass action kinetics. Under mass action kinetics, the intensity function for the kth reaction is

$$\lambda_k(x) = \tilde{c}_k\binom{x}{x-\nu_k^s} = \frac{\tilde{c}_k}{\prod_{\ell=1}^d \nu_{k\ell}^s!}\prod_{\ell=1}^d \frac{x_\ell!}{(x_\ell-\nu_{k\,\ell}^s)!}1_{\{x_\ell \geq 0\}} \stackrel{\text{def}}{=} c_k \prod_{\ell=1}^d \frac{x_\ell!}{(x_\ell-\nu_{k\,\ell}^s)!}1_{\{x_\ell\geq 0\}}, \tag{1.3}$$

where $\tilde{c}_k$ is a positive constant and c_k is defined by the above equation. Mass action kinetics arises by thinking of $\tilde{c}_k\Delta t$ as the approximate probability that a *particular* set of the molecules needed in the kth reaction will react over a time-period of size Δt, and then counting the number of ways such a reaction could happen. Implicit in the assumption of mass action kinetics is that the vessel under consideration is "well-stirred." For ease of notation we will henceforth drop the indicator functions from our representation of mass action kinetics. More general rates will be discussed in the remark at the top of page six.

1.2. *Numerical methods.* There are a number of numerical methods that produce statistically exact sample paths for the model described above. These include the stochastic simulation algorithm, better known as Gillespie's algorithm ([9, 10]), the first reaction method ([9]), and the next reaction method ([1, 8]). All such algorithms perform the same two basic steps multiple times until a sample path is produced over a desired time interval: first, conditioned on the current state of the system the amount of time that passes until the next reaction takes place, Δ, is computed and second the specific reaction that has taken place is found. If, however, $\sum_k \lambda_k(X(t)) \gg 0$ then $\Delta \approx (\sum_k \lambda_k(X(t)))^{-1} \ll 1$ and the time needed to produce a single exact sample path over a time interval can be prohibitive.

The approximate algorithm "tau-leaping" was developed by Dan Gillespie in [11] in an effort to overcome the problem that Δ may be prohibitively small. The basic idea of tau-leaping is to hold the intensity functions fixed over the time interval $[t_n, t_n + h]$ at the values $\lambda_k(X(t_n))$, where $X(t_n)$ is the current state of the system, and, under this assumption, compute the number of times each reaction takes place over this period. As the waiting times for the reactions are exponentially distributed this leads to the following algorithm.

ALGORITHM 1 (Euler tau-leaping). Set $Z(0) = X(0)$, $t_0 = 0$, $n = 0$ and repeat the following until $t_{n+1} > T$.

1. Set $Z(t_{n+1}) = Z(t_n) + \sum_k \mathcal{P}_{k,n}(\lambda_k(Z(t_n))h)\nu_k$, set $t_{n+1} = t_n + h$, and set $n = n + 1$, where $\mathcal{P}_{k,n}(x)$ are independent Poisson random variables with parameters x.

Several improvements and modifications have been made to the basic algorithm described above over the years. However, they are mainly concerned with how to choose the step-size adaptively [4, 12] and/or how to ensure that population values do not go negative during the course of a simulation [2, 3, 5], and are not explicitly relevant to the current discussion.

Similar to (1.2), a path-wise representation of Euler tau-leaping can be given through a random time change of Poisson processes:

$$Z(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(Z \circ \eta(s))ds \right) \nu_k, \tag{1.4}$$

where $\eta(s) = t_n$ if $t_n \leq s < t_{n+1}$ and the Y_k are as before. Noting that $\int_0^{t_{n+1}} \lambda_k(Z \circ \eta(s))ds = \sum_{i=0}^n \lambda_k(Z(t_i))(t_{i+1} - t_i)$ explains our choice to call this method "Euler tau-leaping." Defining the operator

$$(\mathcal{B}_z f)(x) = \sum_k \lambda_k(z)(f(x + \nu_k) - f(x)), \tag{1.5}$$

we see that for $t > 0$

$$\mathbb{E}f(Z(t)) = \mathbb{E}f(Z \circ \eta(t)) + \mathbb{E} \int_{\eta(t)}^t (\mathcal{B}_{Z\circ\eta(t)} f)(Z(s))ds, \tag{1.6}$$

so long as the expectations exist. Further, we note that $\tilde{Z}(t) \stackrel{\text{def}}{=} Z(t) - \sum_k \int_0^t \lambda_k(Z \circ \eta(s))ds\, \nu_k$ is a martingale with quadratic covariation matrix $[\tilde{Z}]_t = \sum_k Y_k \left(\int_0^t \lambda_k(Z \circ \eta(s))ds \right) \nu_k \nu_k^T$.

It is natural to believe that a midpoint type method would be more accurate than an Euler type method in many situations. We therefore define the function

$$\rho(z) \stackrel{\text{def}}{=} z + \frac{1}{2} h \sum_k \lambda_k(z)\nu_k,$$

which computes an approximate midpoint for the system assuming the state of the system is z and the time-step is h.

ALGORITHM 2 (Midpoint tau-leaping). Set $\mathcal{Z}(0) = X(0)$, $t_0 = 0$, $n = 0$ and repeat the following until $t_{n+1} > T$.

1. Set $\mathcal{Z}(t_{n+1}) = \mathcal{Z}(t_n) + \sum_k \mathcal{P}_{k,n}(\lambda_k \circ \rho \circ \mathcal{Z}(t_n)h)\nu_k$, set $t_{n+1} = t_n + h$, and set $n = n + 1$, where $\mathcal{P}_{k,n}(x)$ are independent Poisson random variables with parameters x.

Similar to (1.2) and (1.4), $\mathcal{Z}(t)$ can be represented via a random time change of Poisson processes:

$$\mathcal{Z}(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k \circ \rho \circ \mathcal{Z} \circ \eta(s)ds \right) \nu_k,$$

where $\eta(\cdot)$ is as before. For $\mathcal{B}_z$ defined via (1.5) and any $0 < t$ and any function f

$$\mathbb{E}f(\mathcal{Z}(t)) = \mathbb{E}f(\mathcal{Z}\circ\eta(t)) + \mathbb{E}\int_{\eta(t)}^{t}(\mathcal{B}_{\rho\circ\mathcal{Z}\circ\eta(t)}f)(\mathcal{Z}(s))ds. \tag{1.7}$$

Finally, $\tilde{\mathcal{Z}}(t) \stackrel{\text{def}}{=} \mathcal{Z}(t) - \sum_k \int_0^t \lambda_k \circ \rho \circ \mathcal{Z} \circ \eta(s)ds\,\nu_k$ is a martingale with quadratic covariation matrix $[\tilde{\mathcal{Z}}]_t = \sum_k Y_k\left(\int_0^t \lambda_k \circ \rho \circ \mathcal{Z} \circ \eta(s)ds\right)\nu_k\nu_k^T$. The main goal of this paper is to show that the midpoint tau-leaping algorithm is indeed more accurate than the Euler tau-leaping method under an appropriate, and natural, scaling described in Section 2.

REMARK. Historically the time discretization parameter for tau-leaping has been τ, thus giving the method its name. We choose to break from this tradition and denote our time-step by h so as not to confuse τ with a stopping time.

1.3. *Previous error analyses.* Under the scaling $h \to 0$ Rathinam et al. performed a consistency check of Euler tau-leaping and found that the local truncation error was $\mathcal{O}(h^2)$ for all moments [18]. They also showed that under this same scaling Euler tau-leaping is first order accurate in a weak sense in the special case that the intensity functions λ_k are linear [18]. Li extended these results by showing that as $h \to 0$ Euler tau-leaping has a strong error (in the L^2 norm) of order $1/2$ and a weak error of order one [17], which agree with classical results pertaining to numerical analysis of SDEs driven by Brownian motions (see, for example, [13]).

Under the scaling $h \to 0$ it is readily seen that midpoint tau-leaping is no more accurate than Euler tau-leaping. This follows since midpoint tau-leaping consists of making an $\mathcal{O}(h^2)$ correction to the intensity functions used in Euler tau-leaping. As $h \to 0$, this correction becomes negligible as Poisson processes "ignore" $\mathcal{O}(h^2)$ corrections, and the accuracy of the two methods will be the same.

We simply note that while the analyses performed in [18] and [17] and the argument made in the previous paragraph are technically correct, performing an analysis as $h \to 0$, independent of the rest of the model, is at odds with the useful regime of tau-leaping. That is, tau-leaping would only be used in a regime where $h \gg \Delta$, where Δ is the expected amount of time between reactions, for otherwise an exact method would be performed. Therefore, we should require that

$$h \gg (\sum_k \lambda_k(Z(t)))^{-1} \qquad \text{or} \qquad h\sum_k \lambda_k(Z(t)) \gg 1, \tag{1.8}$$

where $Z(t)$ is the state of the system. In Section 2 we will present a natural scaling for the models under consideration that does satisfy (1.8) and under which we will perform our analysis.

1.4. *Paper outline.* The remainder of the paper is organized as follows. In Section 2 we give some natural assumptions on the models considered in this paper and introduce the scaling under which we perform our analysis. In Section 3 we perform a strong error analysis for both the Euler and midpoint tau-leaping methods and show that midpoint tau-leaping is the more accurate of the two under our scaling. In Section 4 we perform a weak error analysis of the different methods and again conclude that the midpoint method is more accurate. In Section 5 we present numerical examples demonstrating our results.

2. Assumptions on the model.

2.1. *Scalings of the model and the algorithms.* As discussed in the introduction, tau-leaping methods will only be of use if the time-discretization parameter h satisfies $h\sum_k \lambda_k(Z(t)) \gg 1$ while $(\sum_k \lambda_k(Z(t)))^{-1} \ll 1$, where $Z(t)$ is the state of the system at time t. There are a number of ways for the second condition to hold and a modeling choice must be made. We make the following natural assumptions:

(i) The initial abundance of each species scales with V for some $V \gg 1$.

(ii) Each rate constant satisfies $c_k^V = \mathcal{O}(V^{1-\nu_k^s \cdot \vec{1}})$, where $\vec{1} = [1, 1, \ldots, 1]^T$. In particular, $c_k^V = d_k / V^{1-\nu_k^s \cdot \vec{1}}$ for some $d_k > 0$.

We will denote by X^V the normalized process defined as the vector of abundances divided by V, and will denote by λ_k^V the intensity function defined to be mass action kinetics with rate constants c_k^V. This scaling is the so called "classical scaling" and arises naturally by thinking of V as the volume of the vessel in which the reactions are taking place multiplied by Avogadro's number ([14]). In this case X^V gives the concentration of each species in moles per unit volume. To understand the scaling for the rate constants, consider the case of a reaction requiring as input two constituent molecules: S_1 and S_2. Perhaps $S_1 + S_2 \to S_3$. It is reasonable to assume that the probability that a *particular* pair of S_1 and S_2 molecules meet, and react, in a small time interval is inversely proportional to the volume of the vessel. This same type of logic holds for the cases in which more than two molecules are needed for a reaction to take place (i.e. the probability that three particular molecules meet and react is inversely proportional to the volume squared). For the case that only one molecule is needed for a reaction to take place, it is reasonable to assume that the probability of such a reaction taking place in the next small interval of time for a *particular* molecule should not scale with the volume. See also [19], Chapter 6.

Models that satisfy assumptions (i) and (ii) above have an important property that we will detail here and make use of later. Let $x(t)$ denote the solution to the deterministic initial value problem

$$\dot{x}(t) = F(x(t)) \stackrel{\text{def}}{=} \sum_k d_k x(t)^{\nu_k^s} \nu_k, \qquad x(0) = x_0 \in \mathbb{R}^d_{\geq 0}, \tag{2.1}$$

where d_k is defined in assumption (ii) above, and where for any two vectors $u^v \stackrel{\text{def}}{=} u_1^{v_1} \cdots u_d^{v_d}$ and we adopt the convention that $0^0 = 1$. That is, $x(t)$ is the solution to the corresponding deterministically modeled chemical reaction system with mass action kinetics. It was shown in [14, 15] that for any $\epsilon > 0$ and any $T > 0$, if $X^V(0) = x(0) = x_0$, then

$$\lim_{V \to \infty} P\{ \sup_{t \in [0,T]} |X^V(t) - x(t)| \geq \epsilon \} \to 0. \tag{2.2}$$

Denoting λ_k as *deterministic* mass action kinetics with rate constant d_k, it is an exercise to check that for any reaction, i.e. zeroth order, first order, second order, etc., and any $x \in \mathbb{R}^d_{\geq 0}$

$$\lambda_k^V(Vx) = V\lambda_k(x) + \zeta_k^V(x),$$

where ζ_k^V is uniformly bounded in V and is nonzero only if the reaction requires more than one molecule of a particular species as an input. For example, for the second order reaction $S_1 + S_2 \to S_3$ we have

$$\lambda_k^V(Vx) = \frac{d_k}{V}(Vx_1)(Vx_2) = Vd_k x_1 x_2 = V\lambda_k(x),$$

whereas for the second order reaction $2S_1 \to S_3$ we have

$$\lambda_k^V(Vx) = \frac{d_k}{V} Vx_1 (Vx_1 - 1) = Vd_k x_1^2 - d_k x_1 = V\lambda_k(x) + \zeta_k^V(x),$$

with $\zeta_k^V(x) = -d_k x_1$. The term ζ_k^V will have a true V dependence if three or more molecules of a particular species are required as input. We now make the definition $A_k^V(x) \stackrel{\text{def}}{=} \frac{1}{V}\lambda_k^V(Vx)$, and note that for all $x \in \mathbb{R}^d_{\geq 0}$

$$A_k^V(x) = \lambda_k(x) + \frac{1}{V}\zeta_k^V(x), \tag{2.3}$$

and $A_k^V(x) \equiv 0$ if $x \notin \mathbb{R}^d_{\geq 0}$. Manipulating the definition of A_k^V shows that for all $x \in \mathbb{R}^d$

$$\lambda_k^V(Vx) = VA_k^V(x). \tag{2.4}$$

REMARK. The assumption of mass action kinetics is not critical to the analysis carried out in this paper. Instead, what is critical to this particular analysis is that our kinetics satisfies the scaling (2.4) for A_k^V satisfying (2.3) with λ_k sufficiently smooth.

We now choose a discretization parameter for the approximate methods that is dependent upon the assumptions of the model set out above. We let

$$h^V \stackrel{\text{def}}{=} 1/V^\beta, \tag{2.5}$$

where $0 < \beta < 1$. We note that this scaling satisfies the necessary requirements detailed above as

$$\left(\sum_k \lambda_k^V(Vx)\right)^{-1} = \mathcal{O}(V^{-1}) \ll 1$$
$$V^{-\beta}\left(\sum_k \lambda_k^V(Vx)\right) = \mathcal{O}(V^{1-\beta}) \gg 1.$$

With this choice of time-step, we let Z^V and $\mathcal{Z}^V$ denote the processes generated by Euler and midpoint tau-leaping, respectively, normalized by V. We can now state more clearly what the analysis of this paper will entail. We will consider the case of $V \gg 0$ by letting $V \to \infty$ and consider the relationship of the normalized approximate processes Z^V and $\mathcal{Z}^V$ to the original process X^V, normalized similarly. Note that all three processes converge to the solution of (2.1). We will perform both weak and strong error analyses. In the strong error analysis, we will consider L^1 convergence as opposed to the more standard (at least for systems driven by Brownian motions) L^2 convergence. The reason for this is simple: the Itô isometry makes working with the L^2-norm easier in the Brownian motion case, whereas Poisson processes lend themselves naturally to analysis in the L^1-norm.

We remark that it is clear that the choice of scaling laid out in this section and assumed throughout the paper will not explicitly cover all cases of interest. For example, one may choose to use approximation methods when (i) the abundances of only a strict subset of the constituent species are in an $\mathcal{O}(V)$ scaling regime, or (ii) it is the rate constants themselves that are $\mathcal{O}(V)$ while the abundances are $\mathcal{O}(1)$, or (iii) there is a mixture of the previous two cases with potentially more than two natural scales in the system. Our analysis will not be directly applicable to such cases. However, the purpose of this analysis is not to handle every conceivable case. Instead, our purpose is to try and give a more accurate picture of how different tau-leaping methods approximate the exact solution, both strongly and weakly, in at least one plausible setting and we believe that the analysis detailed in this paper achieves this aim. Further, we believe that error analyses conducted under different modeling assumptions can be carried out in similar fashion.

2.2. *Redefining the kinetics.* Before proceeding to the analysis, we allow ourselves one change to the model detailed in the previous section. As we will be considering approximation methods in which changes to the state of the system are determined by Poisson random variables (which can produce arbitrarily large values), there will always be a positive probability that solutions will leave a region in which the scaling detailed above is valid. Multiple options are available to handle such a situation. One option would be to define a stopping time for when the process leaves a predetermined region in which the scaling regime is valid and then only perform the analysis up to that stopping time. Another option, and the one we choose, is to simply modify the kinetics by multiplying by a cutoff function that makes the intensity functions zero outside such a region. This has the added benefit of guaranteeing the existence of all moments of the processes involved. Note that without this truncation or some other additional assumption guaranteeing the existence of the necessary moments, some of the moment estimates that follow may fail; however, the convergence in probability and convergence in distribution results in Theorems 3.10 and 3.17 would still be valid.

Let $\gamma \geq 0$ be C^∞ with compact support $\Omega_\gamma \subset \mathbb{R}^d_{>0}$, with $\gamma(x) = 1$ for all $x \in B_r(x(t))$ for some $r > 0$, where $x(t)$ satisfies (2.1). Now we redefine our intensity functions by setting

$$\lambda_k^V(x) = \gamma(x/V)c_k^V \prod_{\ell=1}^{d} \frac{x_\ell!}{(x_\ell - \nu_{k\ell}^s)!}, \quad \text{for } x \in \mathbb{R}^d, \tag{2.6}$$

where c_k^V still satisfies the scaling detailed in the previous section. It is easy to check that the redefined kinetics still satisfies $\lambda_k^V(Vx) = VA_k^V(x)$, where now $A_k^V(x)$ has also been redefined by multiplication by $\gamma(x)$. Further, the redefined λ_k^V is identical to the previous function on the domain of interest to us. That is, they only differ if the process leaves the scaling regime of interest. For the remainder of the paper we assume our intensity functions are given by (2.6). Finally, we note that for each k we have the existence of an $L_k > 0$ such that

$$\sup_{x \in \mathbb{R}^d,\, |\alpha| < \infty} |D^\alpha A_k^V(x)| \leq L_k. \tag{2.7}$$

3. Strong error analysis for Euler and midpoint tau-leaping. Throughout this section we assume a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$ with $t_n - t_{n-1} = h^V = V^{-\beta}$ for some $0 < \beta < 1$. In Section 3.1 we give some necessary technical results. In Section 3.2 we give bounds for $\sup_{t \leq T} \mathbb{E}|X^V(t) - Z^V(t)|$ and $\sup_{t \leq T} \mathbb{E}|X^V(t) - \mathcal{Z}^V(t)|$ in terms of V, where $X^V(t)$, $Z^V(t)$, and $\mathcal{Z}^V(t)$ are the normalized processes and satisfy the representations

$$X^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V(X^V(s))ds\right)\nu_k \tag{3.1}$$

$$Z^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V(Z^V \circ \eta(s))ds\right)\nu_k \tag{3.2}$$

$$\mathcal{Z}^V(t) = X^V(0) + \frac{1}{V}\sum_k Y_k\left(V\int_0^t A_k^V \circ \rho^V \circ \mathcal{Z}^V \circ \eta(s)ds\right)\nu_k, \tag{3.3}$$

where

$$\rho^V(z) \stackrel{\text{def}}{=} z + \frac{1}{2}V^{-\beta}\sum_k A_k^V(z)\nu_k,$$

and $\eta(s) = t_n$ for $s \in [t_n, t_{n+1})$. In Sections 3.3 and 3.4 we use different couplings of the processes than those above to provide the exact asymptotics of the error processes $X^V - Z^V$ and $X^V - \mathcal{Z}^V$.

3.1. *Preliminaries.* We present some technical, preliminary concepts that will be used ubiquitously throughout the section. For a more thorough reference of the material presented here, see [6], chapter 6. We begin by defining the following filtrations that are generated by the Poisson processes Y_k,

$$\mathcal{F}_{\tilde{u}} \stackrel{\text{def}}{=} \sigma\{Y_k(s_k) \,:\, s_k \leq u_k\}$$
$$\mathcal{F}_u^i \stackrel{\text{def}}{=} \sigma\{Y_k(r), Y_i(s) \,:\, k \neq i,\ s \leq u,\ r \in [0,\infty)\},$$

where $\tilde{u}$ is a multi-index and u is a scalar.

LEMMA 3.1. *Suppose that $X(t)$ satisfies* (1.2) *with non-negative intensity functions λ_k. For $t \geq 0$ and a choice of k,*

$$\tau_k(t) = \int_0^t \lambda_k(X(s))ds \tag{3.4}$$

is an $\{\mathcal{F}_u^k\}$-stopping time.

PROOF. For $u \geq 0$ let $\alpha(u)$ satisfy

$$\int_0^{\alpha(u)} \lambda_k(X(s))ds = u,$$

where we take $\alpha(u) = \infty$ if $\int_0^\infty \lambda_k(X(s))ds < u$. Then $\alpha(u)$ is adapted to $\mathcal{F}_u^k$ and $\{\tau_k(t) \leq u\} = \{t \leq \alpha(u)\} \in \mathcal{F}_u^k$. □

Therefore, if the processes $X(t)$ and $Z(t)$ satisfy (1.2) with non-negative intensity functions $\lambda_{k,1}$ and $\lambda_{k,2}$, respectively, then for $t, s \geq 0$ and a choice of k

(3.5)
$$\mathbb{E}\left|Y_k\left(\int_0^t \lambda_{k,1}(X(r))dr\right) - Y_k\left(\int_0^s \lambda_{k,2}(Z(r))dr\right)\right| = \mathbb{E}\left|\int_0^t \lambda_{k,1}(X(r))dr - \int_0^s \lambda_{k,2}(Z(r))dr\right|,$$

because (i) both the maximum and minimum of two stopping times are stopping times, and (ii) Y_k is monotone.

Similarly to above, one can show that $\tau(t) \stackrel{\text{def}}{=} (\tau_1(t), \tau_2(t), \dots)$, where $\tau_k(t)$ is as in (3.4), is a multi-parameter $\{\mathcal{F}_{\tilde{u}}\}$-stopping time. We now define the filtration

$$\mathcal{G}_t \stackrel{\text{def}}{=} \mathcal{F}_{\tau(t)}$$

and note that by the conditions of Section 2.2 the centered process

$$\tilde{Y}_k\left(\int_0^t \lambda_k(X(s))ds\right) \stackrel{\text{def}}{=} Y_k\left(\int_0^t \lambda_k(X(s))ds\right) - \int_0^t \lambda_k(X(s))ds, \tag{3.6}$$

is a square integrable martingale, with respect to $\mathcal{G}_t$, with quadratic variation $Y_k\left(\int_0^t \lambda_k(X(s))ds\right)$. This fact will be used repeatedly throughout the paper.

3.2. *Bounds on the strong error.* The following theorems give bounds on the errors $\sup_{t\leq T}\mathbb{E}|X^V(t) - Z^V(t)|$ and $\sup_{t\leq T}\mathbb{E}|X^V(t) - \mathcal{Z}^V(t)|$.

THEOREM 3.2. *Let $X^V(t)$ and $Z^V(t)$ satisfy* (3.1) *and* (3.2)*, respectively, for $t \leq T$. Then there exists a constant $C = C(T) > 0$ such that*

$$\sup_{t\leq T}\mathbb{E}|X^V(t) - Z^V(t)| \leq CV^{-\beta} = Ch^V.$$

PROOF. For $t \in [0, T]$ define $E(t) \stackrel{\text{def}}{=} \mathbb{E}|X^V(t) - Z^V(t)|$. Using (3.5) and (2.7)

$$\begin{aligned} E(t) &\leq (\sum_k |\nu_k|L_k)\mathbb{E}\int_0^t \left|X^V(s) - Z^V \circ \eta(s)\right| ds \\ &\leq (\sum_k |\nu_k|L_k)\int_0^t E(s)ds + (\sum_k |\nu_k|L_k)\mathbb{E}\int_0^t \left|Z^V(s) - Z^V \circ \eta(s)\right| ds. \end{aligned}$$

The second term on the right above can be bounded similarly

$$\mathbb{E}\int_0^t \left|Z^V(s) - Z^V \circ \eta(s)\right| ds \leq (\sum_k |\nu_k|L_k)tV^{-\beta},$$

and the result holds via Gronwall's inequality. □

THEOREM 3.3. *Let $X^V(t)$ and $\mathcal{Z}^V(t)$ satisfy* (3.1) *and* (3.3)*, respectively, for $t \le T$. Then there exists a constant $C = C(T) > 0$ such that*

$$\sup_{t\le T} \mathbb{E}|X^V(t) - \mathcal{Z}^V(t)| \le CV^{-\kappa(\beta)}, \quad \textit{where} \quad \kappa(\beta) = \min\left\{\frac{1+\beta}{2}, 2\beta\right\}.$$

Before proving Theorem 3.3 we present some preliminary material. Let $F^V(x) \stackrel{\text{def}}{=} \sum_k A_k^V(x)\nu_k$ and define

$$U^{V,1}(s) \stackrel{\text{def}}{=} \mathcal{Z}^V(s) - \rho^V \circ \mathcal{Z}^V \circ \eta(s) = \mathcal{Z}^V(s) - \mathcal{Z}^V \circ \eta(s) - \frac{1}{2}V^{-\beta}F^V(\mathcal{Z}^V \circ \eta(s))$$

and

$$\tilde{U}^{V,1}(s) \stackrel{\text{def}}{=} (s - \eta(s) - \frac{1}{2}V^{-\beta})F^V(\mathcal{Z}^V \circ \eta(s)).$$

Then

$$U^{V,1}(s) - \tilde{U}^{V,1}(s) = \tilde{\mathcal{Z}}^V(s) - \tilde{\mathcal{Z}}^V \circ \eta(s) + (s - \eta(s))(F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) - F^V(\mathcal{Z}^V \circ \eta(s))), \tag{3.7}$$

where $\tilde{\mathcal{Z}}^V(t) \stackrel{\text{def}}{=} \mathcal{Z}^V(t) - \int_0^t F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds$ is a martingale.

LEMMA 3.4. *For all $0 < \beta < 1$, there exists a $C > 0$ such that*

$$\sup_{s\le\infty} \mathbb{E}|U^{V,1}(s) - \tilde{U}^{V,1}(s)| \le CV^{-\kappa(\beta)}.$$

PROOF. Clearly, the third term on the right of (3.7) is $\mathcal{O}(V^{-2\beta})$ uniformly in s. Thus,

$$\begin{aligned}
\mathbb{E}|U^{V,1}(s) - \tilde{U}^{V,1}(s)| &\le \mathbb{E}|\tilde{\mathcal{Z}}^V(s) - \tilde{\mathcal{Z}}^V \circ \eta(s)| + c_1V^{-2\beta} \\
&\le \left(\frac{1}{V}\sum_k |\nu_k|^2 \mathbb{E}\int_{\eta(s)}^s A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(r))dr\right)^{1/2} + c_1V^{-2\beta} \\
&\le c_2V^{-(1+\beta)/2} + c_1V^{-2\beta},
\end{aligned}$$

for constants c_1 and c_2 which do not depend upon s. □

LEMMA 3.5. *For all $0 < \beta < 1$ and $0 < t$, and for $\alpha \in \{2, 3, 4, \dots\}$*

$$\lim_{V\to\infty} V^{\alpha\beta} \sup_{s\le t} \mathbb{E}[|U^{V,1}(s) - \tilde{U}^{V,1}(s)|^\alpha] = 0.$$

PROOF. The third term on the right of (3.7) is $\mathcal{O}(V^{-2\beta})$, so

$$\begin{aligned}
\mathbb{E}|U^{V,1}(s) - \tilde{U}^{V,1}(s)|^2 &\le C(\mathbb{E}|\tilde{\mathcal{Z}}^V(s) - \tilde{\mathcal{Z}}^V \circ \eta(s)|^2 + V^{-4\beta}) \\
&\le \frac{C}{V}\sum_k |\nu_k|^2 \mathbb{E}\int_{\eta(s)}^s A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(r))dr + CV^{-4\beta} \\
&= \mathcal{O}(V^{-((1+\beta)\wedge 4\beta)}),
\end{aligned}$$

showing the $\alpha = 2$ case.

It is simple to show that $V^{\alpha\beta}\sup_{s\le t}\mathbb{E}[|U^{V,1}(s) - \tilde{U}^{V,1}(s)|^\alpha]$ is uniformly bounded in V for any $\alpha \in \mathbb{Z}_{\ge 0}$. The $\alpha = 2$ case then gives the necessary bounds for the arbitrary α case. □

Note that by Lemmas 3.4 and 3.5

$$\begin{aligned} A_k^V(\mathcal{Z}^V(s)) - A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) &= \nabla A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) \cdot U^{V,1}(s) + \mathcal{O}(V^{-2\beta}) \\ &= \nabla A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) \cdot \tilde{U}^{V,1}(s) + \mathcal{O}(V^{-\kappa(\beta)}) \end{aligned} \tag{3.8}$$

We finally note that for any bounded function g and any $n \geq 0$

$$\int_{t_n}^{t_{n+1}} g(\eta(s))\tilde{U}^{V,1}(s)ds = 0$$

and so for any $t > 0$

$$\int_0^t g(\eta(s))\tilde{U}^{V,1}(s)ds = \frac{1}{8}\left((2t - 2\eta(t) - V^{-\beta})^2 - V^{-2\beta}\right) g(\eta(t))F^V(\mathcal{Z}^V \circ \eta(t)) = \mathcal{O}(V^{-2\beta}). \tag{3.9}$$

PROOF. (Of Theorem 3.3) For $t \leq T$ define $E(t) \stackrel{\text{def}}{=} \mathbb{E}|X^V(t) - \mathcal{Z}^V(t)|$. Letting c_i denote constants

$$\begin{aligned} E(t) &\leq \sum_k |\nu_k| \mathbb{E}\left| \int_0^t A_k^V(X^V(s))ds - \int_0^t A_k^V \circ \rho^V(\mathcal{Z}^V \circ \eta(s))ds \right| \\ &\leq c_1 \int_0^t E(s)ds + \sum_k |\nu_k| \, \mathbb{E}\left| \int_0^t A_k^V(\mathcal{Z}^V(s)) - A_k^V \circ \rho^V(\mathcal{Z}^V \circ \eta(s))ds \right| \\ &\leq c_1 \int_0^t E(s)ds + c_2 V^{-\kappa(\beta)}, \end{aligned}$$

where the final inequality used both (3.8) and (3.9). The result now follows from Gronwall's inequality. □

3.3. *Exact asymptotics for Euler tau-leaping.* Throughout this section and the next all convergences are understood to hold on bounded intervals. More explicitly, we write $X^V \to X$ if $\lim_{V\to\infty} P\{\sup_{t\leq T} |X^V(t) - X(t)| > \epsilon\} = 0$ for all $\epsilon > 0$ and $T > 0$. Because of the simplifying assumptions made on the kinetics in Section 2.2 it is not difficult to show that $X^V \to X$ also implies $\lim_{V\to\infty} \mathbb{E}\sup_{t\leq T} |X^V(t) - X(t)| = 0$. In light of this, when we write $X^V = Z^V + \mathcal{O}(V^{-p})$ for some $p > 0$ in this section and the next we mean that for any $T > 0$ there exists a $C(T)$ such that

$$\lim_{V\to\infty} V^p \mathbb{E}\sup_{t\leq T} |X^V(t) - Z^V(t)| \leq C(T).$$

Finally, recall that $F^V(x) = \sum_k A_k^V(x)\nu_k$ and note that the function $F(x)$ and the deterministic process $x(s)$ used in the characterization of the error processes are defined via (2.1).

Theorem 3.2 suggests that $X^V - Z^V$ scales like $V^{-\beta}$. In this section we make this precise by characterizing the limiting behavior of $V^\beta(X^V - Z^V)$, as $V \to \infty$. To get the exact asymptotics for the Euler tau-leap method we will use the following coupling of the processes involved

$$\begin{aligned} X^V(t) = X^V(0) + \frac{1}{V}\sum_k \Bigg[& Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds\right) \\ & + Y_{k,2}\left(V\int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))ds\right)\Bigg]\nu_k \end{aligned} \tag{3.10}$$

$$\begin{aligned} Z^V(t) = X^V(0) + \frac{1}{V}\sum_k \Bigg[& Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))\, ds\right) \\ & + Y_{k,3}\left(V\int_0^t A_k^V(Z^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(Z^V \circ \eta(s))ds\right)\Bigg]\nu_k. \end{aligned} \tag{3.11}$$

It is important to note that the distributions of X^V and Z^V defined via (3.10) and (3.11) are the same as those for the processes defined via (3.1) and (3.2).

The following Lemma is easy to prove using Doob's inequality.

LEMMA 3.6. *For X^V and Z^V given by* (3.10) *and* (3.11)*, $X^V - Z^V \to 0$.*

Combining Lemma 3.6 and (2.2) shows that $Z^V - x \to 0$, where x is the solution to the associated ODE. Similarly $Z^V \circ \eta - x \to 0$. These facts will be used throughout this section.

Centering the Poisson processes, we have

$$\begin{aligned} X^V(t) - Z^V(t) &= M^V(t) + \int_0^t F^V(X^V(s)) - F^V(Z^V \circ \eta(s))ds \\ &= M^V(t) + \int_0^t F^V(X^V(s)) - F^V(Z^V(s))ds + \int_0^t F^V(Z^V(s)) - F^V(Z^V \circ \eta(s))ds, \end{aligned} \tag{3.12}$$

where M^V is a martingale.

To obtain the desired results, we must understand the behavior of the first and third terms on the right of (3.12). We begin by considering the third term. We begin by defining $U^{V,2}$ and $\tilde{U}^{V,2}$ by

$$U^{V,2}(s) \stackrel{\text{def}}{=} Z^V(s) - Z^V \circ \eta(s), \qquad \tilde{U}^{V,2}(s) \stackrel{\text{def}}{=} (s - \eta(s))F^V(Z^V \circ \eta(s)).$$

Then,

$$U^{V,2}(s) - \tilde{U}^{V,2}(s) = \tilde{Z}^V(s) - \tilde{Z}^V \circ \eta(s),$$

where $\tilde{Z}^V(t) \stackrel{\text{def}}{=} Z^V(t) - \int_0^t F^V(Z^V \circ \eta(s))ds$ is a martingale. Thus

$$\begin{aligned} F^V(Z^V(s)) - F^V(Z^V \circ \eta(s)) &= DF^V(Z^V \circ \eta(s))U^{V,2}(s) + \mathcal{O}(V^{-2\beta}) \\ &= DF^V(Z^V \circ \eta(s))\tilde{U}^{V,2}(s) + DF^V(Z^V \circ \eta(s))(U^{V,2}(s) - \tilde{U}^{V,2}(s)) + \mathcal{O}(V^{-2\beta}). \end{aligned} \tag{3.13}$$

LEMMA 3.7. *For all $0 < \beta < 1$, $0 < t$, and $\alpha \in \{2, 3, 4, \dots\}$*

$$\lim_{V \to \infty} V^{\alpha\beta} \sup_{s \le t} \mathbb{E}[|U^{V,2}(s) - \tilde{U}^{V,2}(s)|^\alpha] = 0.$$

PROOF. The proof is similar to that of Lemma 3.5. □

We may now characterize the limiting behavior of the third term of (3.12).

LEMMA 3.8. *For $0 < \beta < 1$ and any $t > 0$,*

$$V^\beta \int_0^t F^V(Z^V(s)) - F^V(Z^V \circ \eta(s))ds \to \frac{1}{2}\int_0^t DF(x(s))F(x(s))ds.$$

PROOF. By (3.13) and Lemma 3.7

$$V^\beta \int_0^t F^V(Z^V(s)) - F^V(Z^V \circ \eta(s))ds = V^\beta \int_0^t DF^V(Z^V \circ \eta(s))F^V(Z^V \circ \eta(s))(s - \eta(s))ds + \epsilon_1^V(t),$$

where $\epsilon_1^V \to 0$ as $V \to \infty$. By Lemma 3.6 convergence results similar to (2.2) hold for the process $Z^V \circ \eta$, and because $\int_{\eta(s)}^{\eta(s)+V^{-\beta}} (r - \eta(s))dr = \frac{1}{2}V^{-2\beta}$, the lemma holds as stated. □

Turning now to M^V, we observe that the quadratic covariation is

$$[M^V]_t = \frac{1}{V^2}\sum_k (N^V_{k,2}(t) + N^V_{k,3}(t))\nu_k\nu_k^T,$$

where

$$N^V_{k,2}(t) \stackrel{\text{def}}{=} Y_k\Big(V\int_0^t A^V_k(X^V(s)) - A^V_k(X^V(s)) \wedge A^V_k(Z^V\circ\eta(s))\Big)$$
$$N^V_{k,3}(t) \stackrel{\text{def}}{=} Y_k\Big(V\int_0^t A^V_k(Z^V\circ\eta(s)) - A^V_k(X^V(s)) \wedge A^V_k(Z^V\circ\eta(s))\Big),$$

which as $V\to\infty$ is asymptotic to

$$\frac{1}{V}\sum_k \int_0^t |A^V_k(X^V(s)) - A^V_k(Z^V\circ\eta(s))|ds\ \nu_k\nu_k^T. \tag{3.14}$$

We have the following Lemma.

LEMMA 3.9. *For $0<\beta<1$, $V^\beta M^V\to 0$, as $V\to\infty$.*

PROOF. Multiplying (3.12) by V^α, we see that $V^\alpha(X^V - Z^V)\to 0$ provided $\alpha<\beta$ (so that the third term on the right goes to zero) and provided $V^\alpha M^V\to 0$. By the martingale central limit theorem, the latter convergence holds provided $V^{2\alpha}[M^V]\to 0$ (see Lemma A.2 in Appendix A). Let $\alpha_0 = \sup\{\alpha : \alpha\le\beta,\ V^{2\alpha}[M^V]\to 0\}$. Since $\alpha_0<1$, we have that $2\alpha_0 - 1 < \alpha_0\le\beta$, which implies by the definition of α_0 that $V^{2\alpha_0-1}(X^V - Z^V)\to 0$. Therefore

$$\begin{aligned}
V^{2\alpha_0}[M^V]_t &\approx \sum_k\int_0^t V^{2\alpha_0-1}|A^V_k(X^V(s)) - A^V_k(Z^V\circ\eta(s))|ds\ \nu_k\nu_k^T\\
&\approx \sum_k\int_0^t V^{2\alpha_0-1}|\nabla A^V_k(Z^V\circ\eta(s))\cdot(Z^V(s) - Z^V\circ\eta(s))|ds\ \nu_k\nu_k^T\\
&\approx \sum_k\int_0^t V^{2\alpha_0-1}|\nabla A^V_k(Z^V\circ\eta(s))\cdot F^V(Z^V\circ\eta(s))|(s-\eta(s))ds\ \nu_k\nu_k^T,
\end{aligned}$$

where in the second approximation we used that $V^{2\alpha_0-1}(X^V - Z^V)\to 0$, in the third approximation we substituted $\tilde U^{V,2}(s)$ for $U^{V,2}(s)$, and by $f\approx g$ we mean $f-g\to 0$ as $V\to\infty$. The last expression goes to zero whenever $2\alpha_0 - 1<\beta$, hence the convergence holds. □

We now have the following theorem characterizing the behavior of $V^\beta(X^V - Z^V)$.

THEOREM 3.10. *For X^V and Z^V given by* (3.10) *and* (3.11) *and for $0<\beta<1$, $V^\beta(X^V - Z^V)\to\mathcal{E}$, where $\mathcal{E}$ is the solution to*

$$\mathcal{E}(t) = \int_0^t DF(x(s))\mathcal{E}(s)ds + \frac{1}{2}\int_0^t DF(x(s))F(x(s))ds,\qquad \mathcal{E}(0)=0. \tag{3.15}$$

PROOF. Multiply (3.12) by V^β and observe that

$$V^\beta\int_0^t F^V(X^V(s)) - F^V(Z^V(s))ds \approx \int_0^t DF^V(Z^V(s))V^\beta(X^V(s) - Z^V(s))ds.$$

The theorem now follows directly from Lemmas 3.8 and 3.9. □

3.4. *Exact asymptotics for midpoint tau-leaping.* Throughout this section the Hessian matrix associated with a real valued function g will be denoted by Hg. Also, for any vector U, we will denote by $U^T HF^V(x)U$ the vector whose ith component is $U^T HF_i^V U$, and similarly for F.

The goal of this section is to characterize the limiting behavior of $V^{\kappa(\beta)}(X^V(t) - \mathcal{Z}^V(t))$ where

$$\kappa(\beta) = \min\{2\beta, (1+\beta)/2\} = \begin{cases} 2\beta & \beta < 1/3 \\ (1+\beta)/2 & \beta \geq 1/3 \end{cases}.$$

To get the exact asymptotics for the midpoint method we will use the following representation of the processes involved.

$$X^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\, ds\right) + Y_{k,2}\left(V\int_0^t A_k^V(X^V(s)) - A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds\right)\right]\nu_k \tag{3.16}$$

$$\mathcal{Z}^V(t) = X^V(0) + \frac{1}{V}\sum_k \left[Y_{k,1}\left(V\int_0^t A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\, ds\right) + Y_{k,3}\left(V\int_0^t A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) - A_k^V(X^V(s)) \wedge A_k^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds\right)\right]\nu_k. \tag{3.17}$$

The following is similar to Lemma 3.6.

LEMMA 3.11. *For X^V and $\mathcal{Z}^V$ given by* (3.16) *and* (3.17)*, $X^V - \mathcal{Z}^V \to 0$.*

Combining Lemma 3.11 and (2.2) shows that $\mathcal{Z}^V - x \to 0$, where x is the solution to the associated ODE. Similarly $\mathcal{Z}^V \circ \eta - x \to 0$. These facts will be used throughout this section.

Centering the Poisson processes, we have

$$\begin{aligned} X^V(t) - \mathcal{Z}^V(t) &= M^V(t) + \int_0^t F^V(X^V(s)) - F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds \\ &= M^V(t) + \int_0^t F^V(X^V(s)) - F^V(\mathcal{Z}^V(s))ds + \int_0^t F^V(\mathcal{Z}^V(s)) - F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds, \end{aligned} \tag{3.18}$$

where M^V is a martingale.

As before, we must understand the behavior of the first and third terms on the right of (3.18). We begin by considering the third term. Proceeding as in the previous sections, we define $U^{V,3}$ and $\tilde{U}^{V,3}$ as

$$U^{V,3}(s) \stackrel{\text{def}}{=} \mathcal{Z}^V(s) - \rho^V \circ \mathcal{Z}^V \circ \eta(s) = \mathcal{Z}^V(s) - \mathcal{Z}^V \circ \eta(s) - \frac{1}{2}V^{-\beta}F^V(\mathcal{Z}^V \circ \eta(s))$$

and

$$\tilde{U}^{V,3}(s) \stackrel{\text{def}}{=} (s - \eta(s) - \frac{1}{2}V^{-\beta})F^V(\mathcal{Z}^V \circ \eta(s)).$$

Then

$$U^{V,3}(s) - \tilde{U}^{V,3}(s) = \tilde{\mathcal{Z}}^V(s) - \tilde{\mathcal{Z}}^V \circ \eta(s) + (s - \eta(s))(F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) - F^V(\mathcal{Z}^V \circ \eta(s))), \tag{3.19}$$

where $\tilde{\mathcal{Z}}^V(t) \stackrel{\text{def}}{=} \mathcal{Z}^V(t) - \int_0^t F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds$ is a martingale. Then

$$\begin{aligned} F^V(\mathcal{Z}^V(s)) - F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) &= DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))U^{V,3}(s) \\ &\quad + \frac{1}{2}U^{V,3}(s)^T HF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))U^{V,3}(s) + \mathcal{O}(V^{-3\beta}) \\ &= DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\tilde{U}^{V,3}(s) \\ &\quad + \frac{1}{2}U^{V,3}(s)^T HF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))U^{V,3}(s) \\ &\quad + DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))(U^{V,3}(s) - \tilde{U}^{V,3}(s)) + \mathcal{O}(V^{-3\beta}). \end{aligned} \tag{3.20}$$

LEMMA 3.12. *For all $0 < \beta < 1$, $0 < t$, and $\alpha \in \{2, 3, 4, \dots\}$*

$$\lim_{V\to\infty} V^{\alpha\beta} \sup_{s\le t} \mathbb{E}[|U^{V,3}(s) - \tilde{U}^{V,3}(s)|^\alpha] = 0.$$

PROOF. The proof is similar to Lemma 3.5. □

Let

$$\kappa_1(\beta) = \min\{2\beta, \beta + 1/2\} = \begin{cases} 2\beta & \beta < 1/2 \\ \beta + 1/2 & \beta \ge 1/2 \end{cases}.$$

Note that $\kappa_1(\beta) \ge \kappa(\beta)$ for all $\beta \ge 0$.

LEMMA 3.13. *For $0 < \beta < \frac{1}{2}$ and each $t > 0$,*

$$V^{2\beta} \int_0^t DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))(U^{V,3}(s) - \tilde{U}^{V,3}(s))ds \to \frac{1}{4}\int_0^t DF(x(s))^2 F(x(s))ds, \tag{3.21}$$

for $\beta = \frac{1}{2}$

$$V \int_0^t DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))(U^{V,3}(s) - \tilde{U}^{V,3}(s))ds \Rightarrow M_1(t) + \frac{1}{4}\int_0^t DF(x(s))^2 F(x(s))ds, \tag{3.22}$$

and for $\frac{1}{2} < \beta < 1$,

$$V^{\beta+1/2} \int_0^t DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))(U^{V,3}(s) - \tilde{U}^{V,3}(s))ds \Rightarrow M_1(t), \tag{3.23}$$

where M_1 is a mean zero Gaussian process with independent increments and quadratic covariation

$$[M_1]_t = \frac{1}{3}\int_0^t \sum_k A_k(x(s))DF(x(s))\nu_k\nu_k^T DF(x(s))^T ds. \tag{3.24}$$

PROOF. By Lemma A.1 in Appendix A,

$$\begin{aligned} M_1^V(t) \stackrel{\text{def}}{=} &\int_0^t DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))(\tilde{\mathcal{Z}}^V(s) - \tilde{\mathcal{Z}}^V \circ \eta(s))ds \\ &+ DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(t))(\tilde{\mathcal{Z}}^V(t) - \tilde{\mathcal{Z}}^V \circ \eta(t))(\eta(t) + V^{-\beta} - t) \end{aligned}$$

is a martingale and its quadratic covariation matrix is

$$\int_0^t (\eta(s) + V^{-\beta} - s)^2 DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))d[\tilde{\mathcal{Z}}^V]_s DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))^T.$$

Noting that $\int_{\eta(s)}^{\eta(s)+V^{-\beta}}(\eta(r)+V^{-\beta}-r)^2dr=\frac{1}{3}V^{-3\beta}$, it follows that

$$V^{2\beta+1}[M_1^V]_t\to\frac{1}{3}\int_0^t\sum_k A_k(x(s))DF(x(s))\nu_k\nu_k^TDF(x(s))^Tds,$$

so by the martingale central limit theorem $V^{\beta+1/2}M_1^V$ converges in distribution to a mean zero Gaussian process with independent increments and quadratic variation (3.24).

Since $V^{1/2}(\tilde{\mathcal{Z}}^V-\tilde{\mathcal{Z}}^V\circ\eta)\to 0$, the integral on the left of (3.21), (3.22) and (3.23) can be replaced by

$$M_1^V(t)+\int_0^t DF^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))(s-\eta(s))(F^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))-F^V(\mathcal{Z}^V\circ\eta(s)))ds \tag{3.25}$$

without changing the limits. The second term in (3.25) multiplied by $V^{2\beta}$ converges to $\frac{1}{4}\int_0^t DF(x(s))^2F(x(s))ds$ on bounded time intervals and the three limits follow. □

LEMMA 3.14. *For* $0<\beta<1$,

$$V^{2\beta}\frac{1}{2}\int_0^t U^{V,3}(s)^THF^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))U^{V,3}(s)ds\to\frac{1}{24}\int_0^t F(x(s))^THF(x(s))F(x(s))ds.$$

PROOF. By Lemma 3.12 we can replace $U^{V,3}$by $\tilde{U}^{V,3}$. Observing that $\int_{\eta(s)}^{\eta(s)+V^{-\beta}}(s-\eta(s)-\frac{1}{2}V^{-\beta})^2ds=\frac{1}{12}V^{-3\beta}$,

$$V^{2\beta}\frac{1}{2}\int_0^t(s-\eta(s)-\frac{1}{2}V^{-\beta})^2F^V(\mathcal{Z}^V\circ\eta(s))^THF^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))F^V(\mathcal{Z}^V\circ\eta(s))ds$$

converges as claimed. □

We may now characterize the behavior of the third term of (3.18).

LEMMA 3.15. *Let*

$$R^V(t)=\int_0^t(s-\eta(s)-\frac{1}{2}V^{-\beta})DF^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))F^V(\mathcal{Z}\circ\eta(s))ds.$$

Then for $0<\beta<\frac{1}{2}$,

$$\begin{aligned}&V^{2\beta}\left(\int_0^t\left(F^V(\mathcal{Z}^V(s))-F^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))\right)ds-R^V(t)\right)\\&\quad\to\frac{1}{4}\int_0^t DF(x(s))^2F(x(s))ds+\frac{1}{24}\int_0^t F(x(s))^THF(x(s))F(x(s))ds,\end{aligned}$$

for $\beta=\frac{1}{2}$,

$$\begin{aligned}&V\left(\int_0^t\left(F^V(\mathcal{Z}^V(s))-F^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s))\right)ds-R^V(t)\right)\\&\quad\Rightarrow M_1(t)+\frac{1}{4}\int_0^t DF(x(s))^2F(x(s))ds+\frac{1}{24}\int_0^t F(x(s))^THF(x(s))F(x(s))ds\end{aligned}$$

and for $\frac{1}{2}<\beta<1$,

$$V^{\beta+1/2}\int_0^t(F^V(\mathcal{Z}^V(s))-F^V(\rho^V\circ\mathcal{Z}^V\circ\eta(s)))ds\Rightarrow M_1(t).$$

REMARK. Note that $V^{2\beta}R^V$ is uniformly bounded, $R^V \circ \eta \equiv 0$, and

$$R^V(t) = \frac{1}{2}\left[(t-\eta(t))^2 - (t-\eta(t))V^{-\beta}\right] DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(t))F^V(\mathcal{Z} \circ \eta(t)).$$

PROOF. The lemma follows from (3.20), the previous lemmas, and by noting that $\int_0^t DF^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))\tilde{U}^{V,3}(s)ds = R^V(t)$. □

We now turn to M^V and observe that

$$[M^V]_t = \frac{1}{V^2}\sum_k (N^V_{k,2}(t) + N^V_{k,3}(t))\nu_k\nu_k^T,$$

where

$$N^V_{k,2}(t) \stackrel{\text{def}}{=} Y_k\left(V\int_0^t A^V_k(X^V(s)) - A^V_k(X^V(s)) \wedge A^V_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds\right)$$
$$N^V_{k,3}(t) \stackrel{\text{def}}{=} Y_k\left(V\int_0^t A^V_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) - A^V_k(X^V(s)) \wedge A^V_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds\right),$$

which as $V \to \infty$ is asymptotic to

$$\frac{1}{V}\sum_k \int_0^t |A^V_k(X^V(s)) - A^V_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s))|ds\ \nu_k\nu_k^T.$$

Consequently, we have the following.

LEMMA 3.16. *For $0 < \beta < 1$, $V^{(1+\beta)/2}M^V \Rightarrow M$ where M is a mean-zero Gaussian process with independent increments and quadratic covariation*

$$[M]_t = \sum_k \frac{1}{4}\int_0^t |\nabla A_k(x(s)) \cdot F(x(s))|ds\ \nu_k\nu_k^T.$$

PROOF. Multiplying (3.18) by V^α, we see that $V^\alpha(X^V - \mathcal{Z}^V) \to 0$ provided $\alpha < \kappa_1(\beta)$ (so that the third term on the right goes to zero) and provided $V^\alpha M^V \to 0$. By the martingale central limit theorem, the latter convergence holds provided $V^{2\alpha}[M^V] \to 0$. Let $\alpha_0 = \sup\{\alpha : \alpha \le (\beta+1)/2,\ V^{2\alpha}[M^V] \to 0\}$. We make two observations. First, because $\alpha_0 < 1$, we have that $2\alpha_0 - 1 < \alpha_0$. Second, because $\alpha_0 \le (\beta+1)/2$, we have that $2\alpha_0 - 1 \le \beta$, and, in particular, $2\alpha_0 - 1 < \kappa_1(\beta)$ for all $\beta \in (0,1)$. Combining these observations with the definition of α_0 shows that $V^{2(2\alpha_0-1)}[M^V]_t \to 0$ and hence $V^{2\alpha_0-1}(X^V - \mathcal{Z}^V) \to 0$. We now have

$$\begin{aligned} V^{2\alpha_0}[M^V]_t &\approx \sum_k \int_0^t V^{2\alpha_0-1}|A_k(X^V(s)) - A_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s))|ds\ \nu_k\nu_k^T \\ &\approx \sum_k \int_0^t V^{2\alpha_0-1}|s - \eta(s) - \frac{1}{2}V^{-\beta}||\nabla A_k(\rho^V \circ \mathcal{Z}^V \circ \eta(s)) \cdot F^V(\mathcal{Z}^V \circ \eta(s))|ds\ \nu_k\nu_k^T, \end{aligned}$$

where in the second line we used that $V^{2\alpha_0-1}(X^V - Z^V) \to 0$, and then substituted $\tilde{U}^{V,3}(s)$ for $U^{V,3}(s)$. Since the last expression would go to zero if $2\alpha_0 - 1$ were less than β, we see that $2\alpha_0 - 1 = \beta$, that is, $\alpha_0 = (\beta+1)/2$. Furthermore, observing that $\int_{\eta(s)}^{\eta(s)+V^{-\beta}} |s - \eta(s) - \frac{1}{2}V^{-\beta}|ds = \frac{1}{4}V^{-2\beta}$, we see that

$$V^{\beta+1}[M^V]_t = V^{2\alpha_0}[M^V]_t \to \sum_k \frac{1}{4}\int_0^t |\nabla A_k(x(s)) \cdot F(x(s))|ds\ \nu_k\nu_k^T,$$

and the lemma follows by the martingale central limit theorem. □

Collecting the results, we have the following theorem.

THEOREM 3.17. *Let*

$$\mathcal{H}(t) = \frac{1}{6}\int_0^t DF(x(s))^2F(x(s))ds + \frac{1}{24}\int_0^t F(x(s))^T HF(x(s))F(x(s))ds$$

For $0 < \beta < \frac{1}{3}$, $V^{2\beta}(X^V - \mathcal{Z}^V - R^V) \to \mathcal{E}_1$, *where* $\mathcal{E}_1$ *is the solution of*

$$\mathcal{E}_1(t) = \int_0^t DF(x(s))\mathcal{E}_1(s)ds + \mathcal{H}(t), \quad \mathcal{E}_1(0) = 0. \tag{3.26}$$

For $\beta = \frac{1}{3}$, $V^{2\beta}(X^V - \mathcal{Z}^V - R^V) \Rightarrow \mathcal{E}_2$, *where* $\mathcal{E}_2$ *is the solution of*

$$\mathcal{E}_2(t) = M(t) + \int_0^t DF(x(s))\mathcal{E}_2(s)ds + \mathcal{H}(t), \quad \mathcal{E}_2(0) = 0. \tag{3.27}$$

For $\frac{1}{3} < \beta < 1$, $V^{(1+\beta)/2}(X^V - \mathcal{Z}^V) \Rightarrow \mathcal{E}_3$, *where* $\mathcal{E}_3$ *is the solution of*

$$\mathcal{E}_3(t) = M(t) + \int_0^t DF(x(s))\mathcal{E}_3(s)ds, \quad \mathcal{E}_3(0) = 0. \tag{3.28}$$

PROOF. For $\beta \le \frac{1}{3}$, R^V is $\mathcal{O}(V^{-2\beta})$. Subtract R^V from both sides of (3.18) and observe that

$$\int_0^t F^V(X^V(s)) - F^V(\mathcal{Z}^V(s))ds \approx \int_0^t DF^V(\mathcal{Z}^V(s))(X^V(s) - \mathcal{Z}^V(s) - R^V(s))ds + \int_0^t DF^V(\mathcal{Z}^V(s))R^V(s)ds.$$

Since

$$V^{2\beta}\int_0^t DF^V(\mathcal{Z}^V(s))R^V(s)ds \to -\frac{1}{12}\int_0^t DF(x(s))^2F(x(s))ds,$$

the first two parts follow from Lemmas 3.15 and 3.16.

For $\beta > \frac{1}{3}$, $(1+\beta)/2 < 2\beta \wedge \kappa_1(\beta)$, so $V^{(1+\beta)/2}R^V \to 0$ and

$$V^{(1+\beta)/2}\int_0^t F^V(\mathcal{Z}^V(s)) - F^V(\rho^V \circ \mathcal{Z}^V \circ \eta(s))ds \to 0,$$

and the third part follows by Lemma 3.16. □

4. Weak error analysis. As in previous sections, we assume the existence of a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$ with $t_n - t_{n-1} = V^{-\beta}$ for some $0 < \beta < 1$. We also recall that $\eta(s) = t_n$ for $t_n \le s < t_{n+1}$ for each $n \le N-1$.

Let X^V be a Markov process with generator

$$(\mathcal{A}^V f)(x) = \sum_k V A_k^V(x)(f(x + \nu_k/V) - f(x)). \tag{4.1}$$

Defining the operator

$$(\mathcal{B}_z^V f)(x) = \sum_k V A_k^V(z)(f(x + \nu_k/V) - f(x)), \tag{4.2}$$

we suppose that Z^V and $\mathcal{Z}^V$ are processes that satisfy

$$\mathbb{E}f(Z(t)) = \mathbb{E}f(Z\circ\eta(t)) + \mathbb{E}\int_{\eta(t)}^{t}(\mathcal{B}_{Z\circ\eta(t)}f)(Z(s))ds \tag{4.3}$$

and

$$\mathbb{E}f(\mathcal{Z}(t)) = \mathbb{E}f(\mathcal{Z}\circ\eta(t)) + \mathbb{E}\int_{\eta(t)}^{t}(\mathcal{B}_{\rho^V\circ\mathcal{Z}\circ\eta(t)}f)(\mathcal{Z}(s))ds, \tag{4.4}$$

for all $t>0$, respectively.

We begin with the weak error analysis of Euler tau-leaping, which is immediate in light of Theorem 3.10.

THEOREM 4.1. *Let $X^V(t)$ be a Markov process with generator* (4.1) *and let $Z^V(t)$ be a process that satisfies* (4.3) *for the operator* (4.2)*. Then, for any continuously differentiable function f and any $t\le T$*

$$\lim_{V\to\infty} V^{\beta}\left(\mathbb{E}f(X^V(t)) - \mathbb{E}f(Z^V(t))\right) = \mathcal{E}(t)\cdot\nabla f(x(t)),$$

where $\mathcal{E}(t)$ satisfies (3.15).

PROOF. Without loss of generality, we may assume that $X^V(t)$ and $Z^V(t)$ satisfy (3.10) and (3.11), respectively. The proof now follows immediately from a combination of Taylor's theorem and Theorem 3.10. □

REMARK. Because the convergence in Theorem 4.1 is to a constant independent of the step-size of the method, we see that Richardson extrapolation techniques can be carried out. However, we have not given bounds on the next order correction, and so can not say how much more accurate such techniques would be.

We now consider the weak error analysis of the midpoint method.

THEOREM 4.2. *Let $X^V(t)$ be a Markov process with generator* (4.1) *and let $\mathcal{Z}^V(t)$ be a process that satisfies* (4.4) *for the operator* (4.2)*. Then, for any two times continuously differentiable function f with compact support, there exists a constant $C=C(f,T)>0$ such that*

$$V^{2\beta}|\mathbb{E}f(X^V(T)) - \mathbb{E}f(\mathcal{Z}^V(T))| \le C.$$

Before proving Theorem 4.2, some preliminary material is needed. Let $\mathbb{L}^V \stackrel{\text{def}}{=} \{y : y = x/V,\ x\in\mathbb{Z}^d\}$, and for $x\in\mathbb{L}^V$ and a given function f, let

$$v(t,x) = \mathbb{E}_x f(X^V(t)), \tag{4.5}$$

where $\mathbb{E}_x$ represents the expectation conditioned upon $X^V(0)=x$. Standard results give that $v(t,x)$ satisfies the following initial value problem (see, for example, [7] Proposition 1.5)

$$\partial_t v(t,x) = \mathcal{A}^V v(t,x) = \sum_k V A_k^V(x)(v(t,x+\nu_k/V) - v(t,x)), \quad v(0,x)=f(x), \quad x\in\mathbb{L}^V. \tag{4.6}$$

The above equation can be viewed as a linear system by letting x enumerate over $\mathbb{L}^V$ and treating $v(t,x) = v_x(t)$ as functions in time only. It can even be viewed as finite dimensional because of the conditions on the

intensity functions A_k^V. That is, recall that $A_k^V(x) = 0$ for all x outside the bounded set Ω_γ (see Section 2.2); thus, for any such $x \notin \Omega_\gamma$, $v(t,x) = v_x(t) \equiv f(x)$, for all $t > 0$.

For concreteness, we now let M denote the number of reactions for the system under consideration. For $k, \ell \in [1, \ldots, M]$ and $x \in \mathbb{L}^V$ let

$$D_k(t,x) = V(v(t, x + \nu_k/V) - v(t,x)) \tag{4.7}$$

$$D_{k\ell}(t,x) = V(D_k(t, x + \nu_\ell/V) - D_k(t,x)). \tag{4.8}$$

represent approximations to the first and second spatial derivatives of $v(t,x)$, respectively. For notational ease, we have chosen not to explicitly note the V dependence of the functions $v(t,x)$, $D_k(t,x)$, or $D_{k\ell}(t,x)$.

The following lemma, which should be viewed as giving regularity conditions for $v(t,x)$ in the x variable, is instrumental in the proof of Theorem 4.2. The proof is delayed until the end of the section.

LEMMA 4.3. *Let $v(t,x)$, $D_k(t,x)$, and $D_{k\ell}(t,x)$ be given by* (4.5)*,* (4.7)*, and* (4.8)*, respectively, and let $T > 0$. There exists $K_1 > 0$ and $K_2 > 0$ that do not depend upon V such that*

$$\sup_{t \le T} \sup_{k \le M} \sup_{x \in \mathbb{L}^V} |D_k(t,x)| \le K_1 \tag{4.9}$$

$$\sup_{t \le T} \sup_{k,\ell \le M} \sup_{x \in \mathbb{L}^V} |D_{k\ell}(t,x)| \le K_2. \tag{4.10}$$

We will also need the following lemma, which gives regularity conditions for $D_k(t,x)$ in the t variable, and whose proof is also delayed.

LEMMA 4.4. *Let $D_k(t,x)$ be given by* (4.7)*. There exists a $K > 0$ that does not depend upon V such that*

$$\sup_{t \le T} \sup_{k \le M} \sup_{x \in \mathbb{L}^V} |D_k(t+h,x) - D_k(t,x)| \le Kh,$$

for all $h > 0$.

PROOF. **(of Theorem 4.2)**. Define the function $u(t,x) : [0,T] \times \mathbb{L}^V \to \mathbb{R}$ by

$$u(t,x) \stackrel{\text{def}}{=} \mathbb{E}_x f(X^V(T-t)), \tag{4.11}$$

and for any $w(t,x) : \mathbb{R} \times \mathbb{L}^V \to \mathbb{R}$ we define the operator $\mathcal{L}$ by

$$\mathcal{L}w(t,x) \stackrel{\text{def}}{=} \partial_t w(t,x) + \mathcal{A}^V w(t,x) = \partial_t w(t,x) + \sum_k V A_k^V(x)(w(t, x + \nu_k/V) - w(t,x)).$$

Note that $u(t,x) = v(T-t,x)$, where $v(t,x)$ is given by (4.5), and so by (4.6) $\mathcal{L}u(t,x) = 0$ for $t \in [0,T]$ and $x \in \mathbb{L}^V$. We also define the operator

$$\mathcal{L}_z w(t,x) \stackrel{\text{def}}{=} \partial_t w(t,x) + \mathcal{B}_z^V w(t,x) = \partial_t w(t,x) + \sum_k V A_k^V(z)(w(t, x + \nu_k/V) - w(t,x)),$$

so that by virtue of equation (4.4), for $t \le T$ and any differentiable (in t) function $w(t,x)$

$$\mathbb{E}w(t, \mathcal{Z}^V(t)) = \mathbb{E}w(\eta(t), \mathcal{Z}^V \circ \eta(t)) + \int_{\eta(t)}^t \mathbb{E}\mathcal{L}_{\rho^V \circ \mathcal{Z}^V \circ \eta(t)} w(s, \mathcal{Z}^V(s)) ds. \tag{4.12}$$

Recalling (4.11), we see that

$$\mathbb{E}u(T, \mathcal{Z}^V(T)) = \mathbb{E}f(\mathcal{Z}^V(T))$$
$$\mathbb{E}u(T, X^V(T)) = \mathbb{E}u(0, X^V(0)) = \mathbb{E}f(X^V(T)).$$

Therefore by (4.12), and using that $X^V(0) = \mathcal{Z}^V(0)$,

$$\begin{aligned}\mathbb{E}f(\mathcal{Z}^V(T)) - \mathbb{E}f(X^V(T)) &= \mathbb{E}u(T, \mathcal{Z}^V(T)) - \mathbb{E}u(0, \mathcal{Z}^V(0)) \\ &= \sum_{n=0}^{N-1} \mathbb{E}u(t_{n+1}, \mathcal{Z}^V(t_{n+1})) - \mathbb{E}u(t_n, \mathcal{Z}^V(t_n)) \\ &= \sum_{n=0}^{N-1} \mathbb{E}\int_{t_n}^{t_{n+1}} \mathcal{L}_{\rho^V \circ \mathcal{Z}^V(t_n)} u(s, \mathcal{Z}^V(s))ds.\end{aligned}$$

Because $\mathcal{L}u(t,x) \equiv 0$ for $t \le T$ and $x \in \mathbb{L}^V$

$$\begin{aligned}\mathbb{E}\int_{t_n}^{t_{n+1}} \mathcal{L}_{\rho^V \circ \mathcal{Z}^V(t_n)} u(s, \mathcal{Z}^V(s))ds &= \mathbb{E}\int_{t_n}^{t_{n+1}} \mathcal{L}_{\rho^V \circ \mathcal{Z}^V(t_n)} u(s, \mathcal{Z}^V(s)) - \mathcal{L}u(s, \mathcal{Z}^V(s))ds \\ &= \sum_k \mathbb{E}\int_{t_n}^{t_{n+1}} V\left[A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) - A_k^V(\mathcal{Z}^V(s))\right]\left(u(s, \mathcal{Z}^V(s) + \nu_k/V) - u(s, \mathcal{Z}^V(s))\right)ds \\ &= \sum_k \mathbb{E}\int_{t_n}^{t_{n+1}} \left[A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) - A_k^V(\mathcal{Z}^V(s))\right] D_k(T-s, \mathcal{Z}^V(s))ds. \qquad (4.13)\end{aligned}$$

Thus, it is sufficient to prove that each of the integrals in (4.13) are $\mathcal{O}(V^{-3\beta})$. By Lemma 4.4 each integral term in (4.13) can be replaced by

$$I_k^V(t_n) \stackrel{\text{def}}{=} \mathbb{E}\int_{t_n}^{t_{n+1}} \left[A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) - A_k^V(\mathcal{Z}^V(s))\right] D_k(T-t_n, \mathcal{Z}^V(s))ds. \qquad (4.14)$$

The remainder of the proof consists of proving that $I_k^V(t_n) = \mathcal{O}(V^{-3\beta})$.

Letting $g_n^V(x) \stackrel{\text{def}}{=} \left[A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) - A_k^V(x)\right] D_k(T-t_n, x)$ and applying (4.4) to the integrand in (4.14) yields

$$\begin{aligned}I_k^V(t_n) = \mathbb{E}\int_{t_n}^{t_{n+1}} &\left[A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) - A_k^V(\mathcal{Z}^V(t_n))\right] D_k(T-t_n, \mathcal{Z}^V(t_n))ds \\ &+ \sum_j \mathbb{E}\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} V A_j^V(\rho^V \circ \mathcal{Z}^V(t_n))(g_n^V(\mathcal{Z}^V(r) + \nu_j/V) - g^V(\mathcal{Z}^V(r)))drds.\end{aligned}$$

We have $A_k^V(\rho^V \circ \mathcal{Z}^V(t_n)) = A_k^V(\mathcal{Z}^V(t_n)) + \nabla A_k^V(\mathcal{Z}^V(t_n)) \cdot \frac{1}{2}V^{-\beta}\sum_j A_j^V(\mathcal{Z}^V(t_n))\nu_j + \mathcal{O}(V^{-2\beta})$. Thus,

$$I_k^V(t_n) = \sum_j \frac{1}{2}V^{-\beta}\mathbb{E}\int_{t_n}^{t_{n+1}} \nabla A_k^V(\mathcal{Z}^V(t_n)) \cdot \nu_j A_j^V(\mathcal{Z}^V(t_n)) D_k(T-t_n, \mathcal{Z}^V(t_n))ds + \mathcal{O}(V^{-3\beta}) \qquad (4.15)$$

$$+ \sum_j \mathbb{E}\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} V A_j^V(\rho^V \circ \mathcal{Z}^V(t_n))(g_n^V(\mathcal{Z}^V(r) + \nu_j/V) - g^V(\mathcal{Z}^V(r)))drds \qquad (4.16)$$

After some manipulation, the expected value term of (4.16) becomes

$$\mathbb{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} VA_j^V(\rho^V\circ\mathcal{Z}^V(t_n))[A_k^V(\mathcal{Z}^V(r))-A_k^V(\mathcal{Z}^V(r)+\nu_j/V)]D_k(T-t_n,\mathcal{Z}^V(r)+\nu_j/V)dr\,ds$$
$$+\mathbb{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} A_j^V(\rho^V\circ\mathcal{Z}^V(t_n))[A_k^V(\rho^V\circ\mathcal{Z}^V(t_n))-A_k^V(\mathcal{Z}^V(r))]D_{kj}(T-t_n,\mathcal{Z}^V(r))dr\,ds.$$

By Lemma 4.3 the last term above is $\mathcal{O}(V^{-3\beta})$. Taylor's theorem and the fact that $A_j^V(\rho^V\circ\mathcal{Z}^V(t_n)) = A_j^V(\mathcal{Z}^V(t_n)+\mathcal{O}(V^{-\beta})$ then shows us that the expected value term of (4.16) is equal to

$$-\mathbb{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} A_j^V(\mathcal{Z}^V(t_n))\nabla A_k^V(\mathcal{Z}^V(r))\cdot\nu_j D_k(T-t_n,\mathcal{Z}^V(r)+\nu_j/V)dr\,ds+\mathcal{O}(V^{-3\beta})$$
$$=-\mathbb{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} A_j^V(\mathcal{Z}^V(t_n))\nabla A_k^V(\mathcal{Z}^V(r))\cdot\nu_j D_k(T-t_n,\mathcal{Z}^V(r))dr\,ds+\mathcal{O}(V^{-3\beta}), \tag{4.17}$$

where the second equality stems from an application of Lemma 4.3.

By Lemma 4.3, the function $\phi(x) = A_j^V(\mathcal{Z}^V(t_n))\nabla A_k^V(x)\cdot\nu_j D_k(T-t_n,x)$ satisfies $\sup_\ell|\phi(x+\nu_\ell/V)-\phi(x)| = \mathcal{O}(V^{-1})$. Therefore, applying (4.4) to (4.17) shows that (4.17) is equal to

$$-\mathbb{E}\int_{t_n}^{t_{n+1}}\int_{t_n}^{s} A_j^V(\mathcal{Z}^V(t_n))\nabla A_k^V(\mathcal{Z}^V(t_n))\cdot\nu_j D_k(T-t_n,\mathcal{Z}^V(t_n))dr\,ds+\mathcal{O}(V^{-3\beta}).$$

Noting that the sum over j of the above is the negative of (4.15) plus an $\mathcal{O}(V^{-3\beta})$ correction concludes the proof. □

Theorem 4.2 can be strengthened in the case of $\beta < 1/3$.

THEOREM 4.5. *Let $X^V(t)$ be a process with generator* (4.1) *and let $\mathcal{Z}^V(t)$ be a process that satisfies* (4.4) *for the operator* (4.2)*. Suppose also that $\beta < 1/3$. Then, for any continuously differentiable function f,*

$$\lim_{V\to\infty} V^{2\beta}\left(\mathbb{E}f(X^V(T))-\mathbb{E}f(\mathcal{Z}^V(T))\right) = \mathcal{E}_1(T)\cdot\nabla f(x(T)),$$

where $\mathcal{E}_1(t)$ satisfies (3.26)*.*

PROOF. Noting that $R^V(T)\equiv 0$, this is an immediate consequence of Theorem 3.17. □

REMARK. In Theorem 4.1 we provided an explicit asymptotic value for the scaled error of Euler tau-leaping in terms of a solution to a differential equation for all scales, $0<\beta<1$, of the leap step. However, Theorem 4.5 gives a similar result for the midpoint method only in the case $0<\beta<1/3$. For the case $1/3\le\beta<1$, Theorem 4.2 only shows that the error is asymptotically bounded by a constant. The reason for the discrepancy in results is because in Section 3 we were able to show that the dominant component of the pathwise error for Euler tau-leaping for all $\beta\in(0,1)$ and for midpoint tau-leaping for $\beta\in(0,1/3)$ was a term that converged to a deterministic process. However, in the case $\beta\ge 1/3$ for midpoint tau-leaping, the dominant term of the error is a non-zero Gaussian process. We note that this random error process should not be viewed as "extra fluctuations," as they are present in the other cases. In these other cases, they are just dominated by the error that arises from the deterministic "drift" or "bias" of the error process. We leave the exact characterization of the weak error of the midpoint method in the case $\beta\ge 1/3$ as an open problem.

We now present the delayed proofs of Lemma 4.3 and Lemma 4.4.

PROOF. **(Of Lemma 4.3)** Let $C_1 > 0$ be such that

$$\sup_{x\in\mathbb{L}^V}\sup_k |D_k(0,x)| = \sup_{x\in\mathbb{L}^V}\sup_k |V(f(x+\nu_k/V)-f(x))| \le C_1.$$

Using (4.6), a tedious reordering of terms shows that $D_k(t,x)$ satisfies

$$\begin{aligned}\partial_t D_k(t,x) &= \sum_j A_j^V(x)V[D_k(t,x+\nu_j/V)-D_k(t,x)]\\ &\quad+\sum_j (A_j^V(x+\nu_k/V)-A_j^V(x))VD_j(t,x+\nu_k/V).\end{aligned} \tag{4.18}$$

Similarly to viewing $v(t,x)=v_x(t)$ as a finite dimensional linear system, (4.18) can be viewed as a linear system for the variables $D_k(t,x)=D_{\{k,x\}}(t)$, for $k\in[1,\dots,M]$ and $x\in\mathbb{L}^V$. Because $A_j^V(x)\equiv 0$ for all $x\notin\Omega_\gamma$, we see that $\partial_t D_k(t,x)\equiv 0$ for all x such that $x\notin\Omega_\gamma$ and $x+\nu_j/V\notin\Omega_\gamma$ for all $j\in[1,\dots,M]$. Therefore, the system (4.18) can be viewed as finite dimensional also.

Let $\Gamma_1=[1,\dots,M]\times\mathbb{L}^V$. We enumerate the system (4.18) over $b\in\Gamma_1$. That is, for $b=\{k,x\}\in\Gamma_1$ we let $D_b(t)=D_k(t,x)=D_{b_1}(t,b_2)$. After some ordering of the set Γ_1, we let $\mathbb{R}^{\Gamma_1}$ denote the set of (infinite) vectors, v, whose b^{th} component is $v_b\in\mathbb{R}$, and then denote $D(t)\in\mathbb{R}^{\Gamma_1}$ as the vector whose b^{th} component is $D_b(t)$. Next, for each $b=\{k,x\}\in\Gamma_1$, we let

$$S_b \stackrel{\text{def}}{=} \sum_j A_j^V(b_2) = \sum_j A_j^V(x)$$

and let $r_b, R_b\in\mathbb{R}^{\Gamma_1}$ satisfy

$$\begin{aligned}R_b\cdot v &= \sum_j A_j^V(b_2)v_{\{b_1,b_2+\nu_j/V\}}\\ r_b\cdot v &= \sum_j (A_j^V(b_2+\nu_{b_1}/V)-A_j^V(b_2))Vv_{\{j,b_2+\nu_{b_1}/V\}},\end{aligned}$$

for all $v\in\mathbb{R}^{\Gamma_1}$. It is readily seen that for any b both R_b and r_b have at most M non-zero components. Also, by the regularity conditions on the functions A_j^V's, the absolute value of the nonzero terms of r_b are uniformly bounded above by some K, which is independent of V. Finally, note that $R_b\cdot 1 = S_b$. Combining the previous few sentences shows that for any vector $v\in\mathbb{R}^{\Gamma_1}$ we have the two inequalities

$$|R_b\cdot v| = \left|\sum_j A_j^V(b_2)v_{\{b_1,b_2+\nu_j/V\}}\right| \le S_b\|v\|_\infty \tag{4.19}$$

$$|r_b\cdot v| \le KM\|v\|_\infty, \tag{4.20}$$

where $\|v\|_\infty \stackrel{\text{def}}{=} \sup_{b\in\Gamma_1}|v_b|$. We now write (4.18) as

$$D_b'(t) = -VS_bD_b(t) + VR_b\cdot D(t) + r_b\cdot D(t),$$

and so for each $b\in\Gamma_1$

$$\frac{d}{dt}D_b(t)^2 = -2VS_bD_b(t)^2 + 2VD_b(t)R_b\cdot D(t) + 2D_b(t)r_b\cdot D(t). \tag{4.21}$$

Only a finite number of the terms $D_b(t)$ are changing in time and so there is a b^1 and a $t_1\in(0,T]$ for which $|D_{b^1}(t)| = \|D(t)\|_\infty$ for $t\in[0,t_1]$. By (4.19) we have that for this b^1 and any $t\in[0,t_1]$

$$\int_0^t D_{b^1}(s)R_{b^1}\cdot D(s)ds \le \int_0^t S_{b^1}|D_{b^1}(s)|\|D(s)\|_\infty ds = \int_0^t S_{b^1}D_{b^1}(s)^2 ds,$$

which, after integrating (4.21), yields

$$\|D(t)\|_\infty^2 = D_{b^1}(t)^2 \le D_{b^1}(0)^2 + 2\int_0^t D_{b^1}(s) r_{b^1}\cdot D(s)ds \le \|D(0)\|_\infty^2 + 2KM\int_0^t \|D(s)\|_\infty^2 ds,$$

where the final inequality makes use of (4.20). An application of Gronwall's inequality now gives us that for $t \in [0, t_1]$

$$\|D(t)\|_\infty^2 \le \|D(0)\|_\infty^2 e^{2KMt}.$$

To complete the proof, continue this process for $i \ge 2$ by choosing the b^i for which $|D_{b^i}(t)|$ is maximal on the time interval $t_i - t_{i-1}$. We must have $\lim_{i\to\infty} t_i = T$ because (i) there are a finite number of time varying $D_b(t)$'s and (ii) each $D_b(t)$ is differentiable. After taking square roots we find $\sup_{t\le T}\|D(t)\|_\infty \le \|D(0)\|_\infty e^{KMT} \le C_1 e^{KMT}$, which is equivalent to (4.9).

We now turn our attention to showing (4.10), which we show in a similar manner. There is a $C_2 > 0$ such that for all $x \in \mathbb{L}^V$ and $k, \ell \in [1, \dots, M]$,

$$|D_{k\ell}(0,x)| = V^2|f(x+\nu_\ell/V+\nu_k/V) - f(x+\nu_\ell/V) - f(x+\nu_k/V) + f(x)| \le C_2.$$

Another tedious reordering of terms, which makes use of (4.18), shows that $D_{k\ell}(t,x)$ satisfies

$$\begin{aligned}\partial_t D_{k\ell}(t,x) = &\sum_j A_j^V(x)V[D_{k\ell}(t,x+\nu_j/V) - D_{k\ell}(t,x)] + \sum_j (A_j^V(x+\nu_\ell/V) - A_j^V(x))VD_{kj}(t,x+\nu_\ell/V)\\ &+ \sum_j (A_j^V(x+\nu_k/V) - A_j^V(x))VD_{j\ell}(t,x+\nu_k/V) + g_{k\ell}(t,x),\end{aligned}$$

where

$$\begin{aligned}g_{k\ell}(t,x) \stackrel{\text{def}}{=} \sum_j V^2[&A_j^V(x+\nu_\ell/V+\nu_k/V) - A_j^V(x+\nu_\ell/V) - A_j^V(x+\nu_k/V) + A_j^V(x)]\\ &\times D_j(t, x+\nu_\ell/V+\nu_k/V).\end{aligned}$$

By (i) the fact that the second derivative of A_j^V is uniformly (in j and x) bounded and (ii) the bound (4.9), the absolute value of the last term is uniformly (in $t \le T$, x, k, and ℓ) bounded by some $C_3 > 0$.

As we did for both $v(t,x)$ and $D_k(t,x)$, we change perspective by viewing the above as a linear system with state space $\{k,\ell,x\} \in [1,\dots,M]\times[1,\dots,M]\times\mathbb{L}^V = \Gamma_2$, where we again put an ordering on Γ_2 and consider $\mathbb{R}^{\Gamma_2}$ defined similarly to $\mathbb{R}^{\Gamma_1}$. Also similarly to before, we note that only a finite number of the $D_{k,\ell}(t,x)$ are changing in time. For $b = \{k,\ell,x\} \in \Gamma_2$, we see that $D_b(t)$ satisfies

$$D_b'(t) = -S_b V D_b(t) + VR_b\cdot D(t) + r_b\cdot D(t) + g_b(t), \tag{4.22}$$

where $D_b(t), D(t), S_b, R_b$, and r_b are defined similarly as before and where we retain the necessary inequalities: for $v \in \mathbb{R}^{\Gamma_2}$

$$\begin{aligned}|R_b\cdot v| &= \left|\sum_j A_j^V(b_3) v_{\{b_1,b_2,b_3+\nu_j/V\}}\right| \le S_b\|v\|_\infty\\ |r_b\cdot v| &\le 2KM\|v\|_\infty.\end{aligned} \tag{4.23}$$

The rest of the proof is similar to the proof that the $D_k(t,x)$ are uniformly bounded. There is a $b^1 \in \Gamma_2$ and a $t_1 \in (0,T]$ for which $|D_{b^1}(t)| = \|D(t)\|_\infty$ for all $t \in [0,t_1]$. Taking the derivative of $D_{b^1}(t)^2$ while using

(4.22), integrating, and using the bounds (4.23), we have that for this b^1 and any $t \in [0, t_1]$

$$\begin{aligned} D_{b^1}(t)^2 &= D_{b^1}(0)^2 + 2\int_0^t g_{b^1}(s) D_{b^1}(s) ds - 2\int_0^t S_{b^1} V D_{b^1}(s)^2 ds \\ &\quad + 2\int_0^t V R_{b^1} \cdot D(s) D_{b^1}(s) ds + 2\int_0^t r_{b^1} \cdot D(s) D_{b^1}(s) ds \\ &\leq D_b(0)^2 + 2C_3 t + (4KM + 2C_3)\int_0^t \|D(s)\|_\infty^2 ds, \end{aligned}$$

where we used the inequality $x \leq 1 + x^2$ on the term $D_{b^1}(s)$ in the first integral above. Therefore, for $t \leq t_1$

$$\|D(t)\|_\infty^2 \leq \|D(0)\|_\infty^2 + 2C_3 t + (4KM + 2C_3)\int_0^t \|D(s)\|_\infty^2 ds.$$

We continue now by choosing a $b_2 \in \Gamma_2$ such that $|D_{b_2}(t)| = \|D(t)\|_\infty$ for all $t \in [t_1, t_2)$, with $t_1 < t_2 \leq T$. By similar arguments as above we have that for $t \in [t_1, t_2]$

$$\begin{aligned} \|D(t)\|^2 &\leq \|D(t_1)\|_\infty^2 + 2C_3(t - t_1) + (4KM + 2C_3)\int_{t_1}^t \|D(s)\|_\infty^2 ds \\ &\leq \|h(0)\|_\infty^2 + 2C_3 t + (4KM + 2C_3)\int_0^t \|h(s)\|_\infty^2 ds. \end{aligned}$$

Continuing in this manner shows that the above inequality holds for all $t \in [0, T]$ and so a Gronwall inequality gives us that for all $t \leq T$

$$\|D(t)\|_\infty^2 \leq \left(\|D(0)\|_\infty^2 + \frac{2C_3}{4KM + 2C_3} \right) e^{(4KM + 2C_3)T},$$

which, after taking square roots, is equivalent to (4.10). □

PROOF. **(Of Lemma 4.4)** By (4.18), we have that for any $k \in [1, \ldots, M]$ and $x \in \mathbb{L}^V$

$$D_k(t, x) = D_k(0, x) + \sum_j A_j^V(x) \int_0^t D_{kj}(s, x) ds + \sum_j (A_j^V(x + \nu_k/V) - A_j^V(x)) V \int_0^t D_j(s, x + \nu_k/v) ds.$$

The proof is now immediate in light of Lemma 4.3. □

5. Examples.

EXAMPLE 5.1. Consider the case of an irreversible isomerization of one molecule into another. We denote by A the molecule undergoing the isomerization and B the target molecule. We assume that the rate constant associated with this reaction is 1. The pictorial representation for this system is simply

$$A \xrightarrow{1} B.$$

Letting $X(t)$ denote the number of A molecules at time $t \geq 0$, $X(t)$ satisfies

$$X(t) = X(0) - Y\left(\int_0^t X(s) ds \right).$$

Supposing that we start with $V = 10,000$ molecules, we approximate the distribution of $X(1)$ using $200,000$ sample paths constructed using the Gillespie algorithm, which produces statistically exact sample

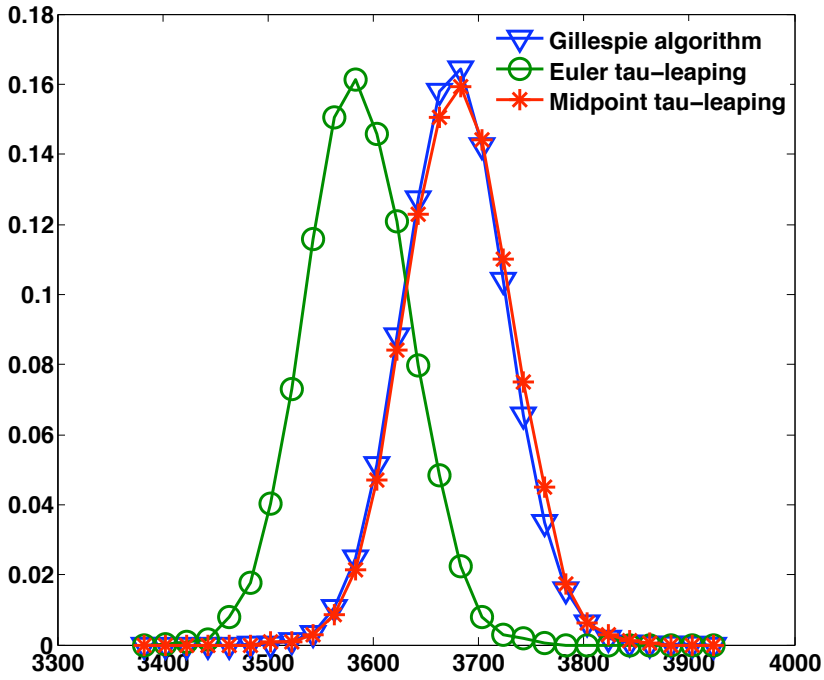

FIG 5.1. *Relative frequency of* $X(1)$ *from* $200,000$ *sample paths constructed using* (i) *Gillespie's algorithm, blue line* ∇ *marker,* (ii) *Euler tau-leaping, green line,* O *marker, and* (iii) *midpoint tau-leaping, red line,* $*$ *marker. The approximated distribution generated via midpoint tau-leaping is clearly closer to the exact distribution than that of Euler tau-leaping.*

paths, Euler tau-leaping with a step-size of $1/20$, and midpoint tau-leaping with a step-size of $1/20$. Note that in this case $1/20 = 1/V^{.325}$, and so $\beta = .325$. The computational results are presented in Figure 5.1, which demonstrate the stronger convergence rate of midpoint tau-leaping as compared to Euler tau-leaping.

It is simple to show that $X(1)$ is a binomial(n,p) random variable with parameters $n = 10,000$ and $p = 1/e$. Therefore, $\mathbb{E}X(1) = 10000/e \approx 3678.8$. The estimated means produced from the 200,000 sample paths of Euler tau-leaping and midpoint tau-leaping were 3585.4 and 3681.4, respectively. Solving for $\mathcal{E}(t)$ of (3.15) for this example yields $\mathcal{E}(t) = (1/2)e^{-t}t$. Theorem 4.1 therefore estimates that Euler tau-leaping should produce a mean $(1/2)e^{-1}10000^{1-.325} \approx 92.2$ smaller than the actual mean, which is in agreement with $3678.8 - 3585.4 = 93.4$. Solving for $\mathcal{E}_1(t)$ of (3.26) for this example yields $\mathcal{E}_1(t) = (1/6)te^{-t}$. Theorem 4.5 therefore estimates that midpoint tau-leaping should produce a mean $(1/6)e^{-1}10000^{1-2*.325} = 4.62$ smaller than the actual mean, which is in agreement with $3678.8 - 3681.4 = -2.6$.

EXAMPLE 5.2. We now consider a simple Lotka-Volterra predator-prey model. Letting A and B represent the prey and predators, respectively, in a given environment we suppose (i) prey reproduce at a certain rate, (ii) interactions between predators and prey benefit the predator while hurting the prey, and (iii) predators die at a certain rate. One possible model for this system is

$$A \xrightarrow{2} 2A, \qquad A + B \xrightarrow{.002} 2B, \qquad B \xrightarrow{2} \emptyset,$$

where a choice of rate constants has been made. Letting $X(t) \in \mathbb{Z}^2_{\geq 0}$ be such that $X_1(t)$ and $X_2(t)$ represent the numbers of prey and predators at time $t > 0$, respectively, $X(t)$ satisfies

(5.1)
$$X(t) = X(0) + Y_1\left(\int_0^t 2X_1(s)ds\right)\begin{bmatrix}1\\0\end{bmatrix} + Y_2\left(\int_0^t .002X_1(s)X_2(s)ds\right)\begin{bmatrix}-1\\1\end{bmatrix} + Y_3\left(\int_0^t 2X_2(s)ds\right)\begin{bmatrix}0\\-1\end{bmatrix}.$$

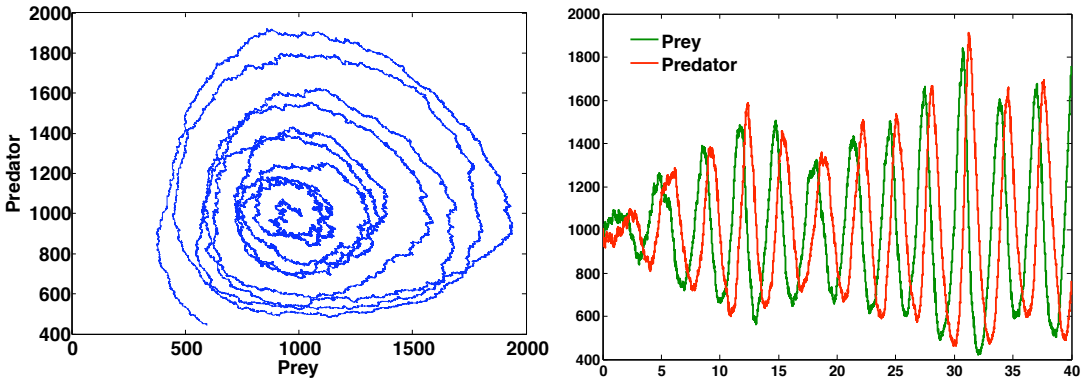

FIG 5.2. *Oscillations in a predator-prey model. In the left image we see the numbers of predators versus the number of prey for a single realization of the system 5.1. In the right image we see the time-series of the numbers of predators and prey for a single realization of 5.1.*

.

We take $X(0) = [1000, 1000]^T$, and so $V = 1000$ for our model. Lotka-Volterra models are famous for producing periodic solutions; this behavior is demonstrated in Figure 5.2.

We approximate the distribution of $X_2(10)$ using $30,000$ sample paths constructed using the Gillespie algorithm, Euler tau-leaping with a step-size of $1/20$, and midpoint tau-leaping with a step-size of $1/20$. Note that in this case $1/20 = 1/V^{.434}$, and so $\beta = .434$. The computational results are presented in Figure 5.3, which again demonstrate the stronger convergence rate of midpoint tau-leaping as compared to Euler tau-leaping.

APPENDIX A

LEMMA A.1. *Let M be a $\{\mathcal{F}_t\}$-martingale, R be bounded and $\{\mathcal{F}_t\}$-adapted, and let $h > 0$. Then for $\eta(t) \equiv [t/h]h$,*

$$\hat{M}(t) = \int_0^t R \circ \eta(s)(M(s) - M \circ \eta(s))ds + R \circ \eta(t)(M(t) - M \circ \eta(t))(\eta(t) + h - t)$$

is an $\{\mathcal{F}_t\}$-martingale and

$$[\hat{M}]_t = \int_0^t (R \circ \eta(r))^2(\eta(r) + h - r)^2 d[M]_r. \tag{A.1}$$

If M is $\mathbb{R}^d$-valued and R is $\mathbb{M}^{m\times d}$-valued, then the quadratic covariation matrix is

$$[\hat{M}]_t = \int_0^t (\eta(r) + h - r)^2 R \circ \eta(r))d[M]_r R^T \circ \eta(r).$$

PROOF. For $t < T - h$,

$$\begin{aligned}
\mathbb{E}[\hat{M}(T)|\mathcal{F}_t] &= \mathbb{E}[\int_0^T R \circ \eta(s)(M(s) - M \circ \eta(s))ds|\mathcal{F}_t] \\
&= \mathbb{E}[\int_0^{\eta(t)+h} R \circ \eta(s)(M(s) - M \circ \eta(s))ds|\mathcal{F}_t] \\
&= \int_0^t R \circ \eta(s)(M(s) - M \circ \eta(s))ds + R \circ \eta(t)(M(t) - M \circ \eta(t))(\eta(t) + h - t).
\end{aligned}$$

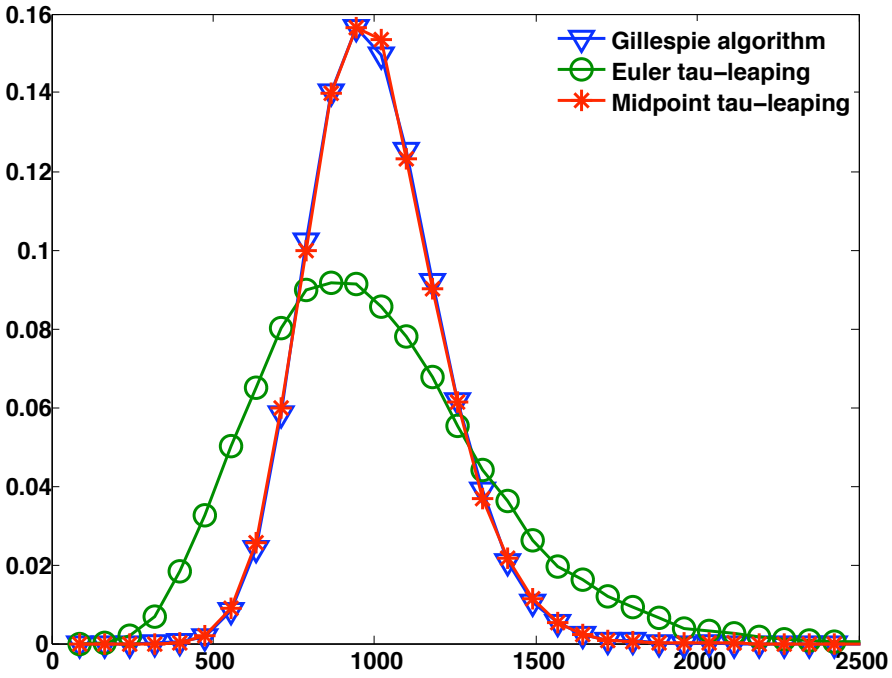

FIG 5.3. *Relative frequency of* $X_2(10)$ *from* $30,000$ *sample paths constructed using* (i) *Gillespie's algorithm, blue line* ∇ *marker,* (ii) *Euler tau-leaping, green line,* O *marker, and* (iii) *midpoint tau-leaping, red line,* $*$ *marker. The approximated distribution generated via midpoint tau-leaping is clearly closer to the exact distribution than that of Euler tau-leaping.*

The case of $T - h \leq t < T$ is similar. $[\hat{M}]$ is just the quadratic variation of the second term on the right, and noting that $\hat{M}$ is continuous at $t = kh$ for all $k = 0, 1, 2 \ldots$, (A.1) follows. □

For completeness we include a statement of the martingale central limit theorem (see [6] for more details).

LEMMA A.2. *Let* $\{M_n\}$ *be a sequence of* $\mathbb{R}^d$*-valued martingales with* $M_n(0) = 0$*. Suppose*

$$\lim_{n\to\infty} \mathbb{E}[\sup_{s\leq t} |M_n(s) - M_n(s-)|] = 0$$

and

$$[M_n^i, M_n^j]_t \to c_{i,j}(t)$$

for all $t > 0$ *where* $C = ((c_{i,j}))$ *is deterministic and continuous. Then* $M_n \Rightarrow M$*, where* M *is Gaussian with independent increments and* $\mathbb{E}[M(t)M(t)^T] = C(t)$.

REFERENCES

[1] David F. Anderson, *A modified next reaction method for simulating chemical systems with time dependent propensities and delays*, J. Chem. Phys. **127** (2007), no. 21, 214107.

[2] ———, *Incorporating postleap checks in tau-leaping*, J. Chem. Phys. **128** (2008), no. 5, 054103.

[3] Yang Cao, Daniel T. Gillespie, and Linda R. Petzold, *Avoiding negative populations in explicit poisson tau-leaping*, J. Chem. Phys. **123** (2005), 054104.

[4] ———, *Efficient step size selection for the tau-leaping simulation method*, J. Chem. Phys. **124** (2006), 044109.

[5] Abhijit Chatterjee and Dionisios G. Vlachos, *Binomial distribution based* τ*-leap accelerated stochastic simulation*, J. Chem. Phys. **122** (2005), 024112.

[6] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: Characterization and convergence*, John Wiley & Sons, New York, 1986.

[7] ———, *Markov processes: Characterization and convergence*, 2 ed., John Wiley & Sons, New York, 2005.

[8] M.A. Gibson and J. Bruck, *Efficient exact stochastic simulation of chemical systems with many species and many channels*, J. Phys. Chem. A **105** (2000), 1876–1889.

[9] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comput. Phys. **22** (1976), 403–434.

[10] ———, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem. **81** (1977), no. 25, 2340–2361.

[11] ———, *Approximate accelerated simulation of chemically reaction systems*, J. Chem. Phys. **115** (2001), no. 4, 1716–1733.

[12] D. T. Gillespie and Linda R. Petzold, *Improved leap-size selection for accelerated stochastic simulation*, J. Chem. Phys. **119** (2003), no. 16, 8229–8234.

[13] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992. MR1214374 (94b:60069)

[14] Thomas G. Kurtz, *The relationship between stochastic and deterministic models for chemical reactions*, J. Chem. Phys. **57** (1972), no. 7, 2976–2978.

[15] ———, *Strong approximation theorems for density dependent Markov chains*, Stoch. Proc. Appl. **6** (1977/78), 223–240.

[16] ———, *Approximation of population processes*, CBMS-NSF Reg. Conf. Series in Appl. Math.: 36, SIAM, 1981.

[17] Tiejun Li, *Analysis of explicit tau-leaping schemes for simulating chemically reacting systems*, SIAM Multiscale Model. Simul. **6** (2007), no. 2, 417–436.

[18] Muruhan Rathinam, Linda R. Petzold, Yang Cao, , and Daniel T. Gillespie, *Consistency and stability of tau-leaping schemes for chemical reaction systems*, SIAM Multiscale Model. Simul. **3** (2005), 867–895.

[19] D. J. Wilkinson, *Stochastic modelling for systems biology*, Chapman and Hall/CRC Press, 2006.

D. F. ANDERSON,
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706
USA
E-MAIL: anderson@math.wisc.edu
URL: http://www.math.wisc.edu/∼anderson/

A. GANGULY,
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706
USA
E-MAIL: ganguly@math.wisc.edu
URL: http://www.math.wisc.edu/∼ganguly/

T. G. KURTZ,
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706
USA
E-MAIL: kurtz@math.wisc.edu
URL: http://www.math.wisc.edu/∼kurtz/